

XX Milk-derived peptides that stimulate Bifidobacterium bifidus
PT Claim 2; Page 3; 25pp; German.
PS
XX This invention describes milk-derived bifidogenic peptides and their
CC active derivatives or fragments, and combinations of them produced by
CC chemical coupling. Such are produced from bovine or human milk by
CC treatment for 2 hr with proteases, then centrifuging to remove fat and
CC acidifying to pH 2 to precipitate proteins. The solution phase is then
CC subjected to reverse-phase HPLC, the fractions adjusted to salt content below
CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by
CC growing Bifidobacterium bifidus and Escherichia coli in presence of the
CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15
CC are selected where Bw = germ count after 16 hr culture of B. bifidus in
CC 50% Elliker broth containing peptide at 0.2 mg/ml, B0 = germ count under
CC similar conditions in a peptide-free control. Ew = germ count after 16
CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at
CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free
CC control. The peptides AAR93865-W93888 are used to treat microbe-related
CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,
CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,
CC microbially induced tumours or degenerative disorders, diarrhoea, colic,
CC abnormalities in oral, intestinal or vaginal microflora, or caries.
XX Sequence 55 AA;
SQ
Query Match 100.0%; Score 305; DB 20; Length 55;
Best Local Similarity 100.0%; Pred. No. 5.7e-30;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 55
DB 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 55
RESULT 2
AAR93351
ID AAR93351 standard; Protein; 182 AA.
AC AAR93351;
XX 25-MAR-2003 (updated)
DT 26-JAN-1994 (first entry)
XX Recombinant human kappa casein.
DE Casein; supplement; milk; pharmaceutical; ss.
XX Homo sapiens.
OS W09315196-A1.
PN 05-AUG-1993.
XX 25-JAN-1993; 93WO-DK00024.
PF 23-JAN-1992; 92DK-0000088.
PR (SYMB-) SYMBICOM AB.
PA Bergstroem S, Hansson L, Hernell O, Stroemqvist M;
XX Toernell J;
PI WPI; 1993-258675/32.
DR N-PSDB; AAR946850.
XX DNA encoding human kappa-casein - used for obtaining recombinant
PT polypeptide(s) for use as nutrient supplements, partic. in infant
PT formulae
PS Claim 44; Page 88-89; 110pp; English.

XX The recombinant human kappa casein is produced in high yields by
CC means of either a eukaryotic or prokaryotic expression system. It
CC is used as a nutrient supplement in milk based products to provide a
CC substantial improvement of the nutritional and biological value of
CC the formulae, making it closer in similarity to human milk. It can
CC also be used as a pharmaceutical.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 182 AA;
SQ
Query Match 100.0%; Score 305; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 55
DB 63 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 117
RESULT 3
AAR72696
ID AAR72696 standard; Protein; 182 AA.
XX AAR72696;
XX 06-NOV-1995 (first entry)
DT Human kappa-casein.
DE Kappa-casein; milk protein.
XX Homo sapiens.
OS US5391497-A.
PN 21-FEB-1995.
XX 13-OCT-1992; 92US-0962569.
PF 13-OCT-1992; 92US-0962569.
PR (COLS) UNIV COLORADO FOUND INC.
PA Chang Y, Ham RG, Jeffers KF, Menon RS;
XX WPI; 1995-160470/21.
DR N-PSDB; AAR89598.
XX DNA encoding human kappa-casein - used for the prodn. of large amts. of
PT highly purified kappa-casein milk protein for infant use.
XX Disclosure; Column 13-16; 14pp; English.
XX A commercial CDNA library prepd. in lambda gtl1 from mRNA obtd. from
CC human breast tissue removed during the third trimester of pregnancy
CC was screened with rabbit anti-bovine kappa-casein cDNA. The cDNA
CC insert of a recombinant phage was amplified by PCR to obtain a
CC full-length clone (AAR89598) encoding human kappa-casein (AAR72696).
XX Sequence 182 AA;
SQ
Query Match 100.0%; Score 305; DB 16; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 55
DB 63 YORRPAIAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPTVVRPNLHPSF 117
RESULT 4
AAR92150

II AAR92150 standard; Protein; 182 AA.
AC AAR92150;
DT 24-OCT-1996 (first entry)
DE Human milk kappa-casein, inhibits rotaviral infection.
XX Milk protein; kappa-casein; rotavirus inhibition; gastroenteritis;
KW infant formulation; immunodeficiency; diarrhoea.
XX Homo sapiens.
XX WO9608269-A1.
XX 21-MAR-1996.
XX 05-MAY-1995; 95WO-US05676.
XX 16-SEP-1994; 94US-0308883.
XX 16-SEP-1994; 94US-0308882.
XX (ABBO) ABBOTT LAB.
XX Baxter JH, Cummings RD, Mukerji P, Prieto PA, Seo AE;
XX WPI; 1996-179724/18.
XX Kappa-casein used as anti-rotaviral infection agent in nutritional
PT product - to prevent, retard or treat rotavirus infection,
PT especially in infants, children and immuno-deficient patients
XX Claim 1; Fig 9; 42pp; English.
CC AAR92150 is human milk kappa-casein which is useful as a component
CC of an enteral nutritional product. Kappa-casein has anti-rotaviral
CC activity, it inhibits the attachment of human rotavirus to mammalian
CC cells. The kappa-casein is useful in the treatment and prevention of
CC onset of gastroenteritis and other diarrhoeal diseases caused by
CC rotaviridae esp. in infants, children or immunodeficient patients.
CC Kappa-casein appeared to have no side effects on the subjects treated
CC and is unlikely to be toxic or cause an allergic reaction.
XX Sequence 182 AA;
SQ
Query Match 100.0%; Score 305; DB 17; Length 182;
Best Local Similarity 100.0%; Pred. No. 2.3e-29;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YORRPAIANNPYPTTYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 55
DB 63 YORRPAIANNPYPTTYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 117
RESULT 5
AAE18216
ID AAE18216 standard; Protein; 181 AA.
XX AAE18216;
AC AAE18216;
XX
DT 07-MAY-2002 (first entry)
XX Human MOL6 protein.
XX Secreted molecule; MOL6 protein; MOLX; cardiomyopathy; atherosclerosis;
KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
KW HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;
KW haematopoietic; thrombolytic; thrombocytopaenia; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; muscular disease; stress;
KW ocular disease; growth disorder; depression; epilepsy; contraceptive;

KW vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;
KW analgesic; vasotropic; hypotensive; gene therapy.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Peptide 1..24 /label= Signal_peptide
FT Protein 25..181 /label= Mature_MOL6_protein
XX
XX WO200206339-A2.

24-JAN-2002.

03-JUL-2001; 2001WO-US21249.

03-JUL-2000; 2000US-215854P.

03-JUL-2000; 2000US-215856P.

03-JUL-2000; 2000US-215902P.

07-JUL-2000; 2000US-216585P.

07-JUL-2000; 2000US-216586P.

17-JUL-2000; 2000US-216722P.

17-JUL-2000; 2000US-218622P.

17-JUL-2000; 2000US-218992P.

14-FEB-2001; 2001US-268734P.

(CURA-) CURAGEN CORP.

Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;

Ratturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;

Alsbrook J, Lepley DM, Shen L, Burgess CE, Shinkles RA;

Padigaru M;

WPI; 2002-155038/20.

N-PSDB; AAD28950.

Nucleic acids encoding secreted polypeptides, designated MOLX

polypeptides, useful for treating a MOLX-associated disorder, e.g.

cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -

Claim 1; Page 78; 223pp; English.

The patent discloses nucleic acid sequences encoding novel secreted

molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL

protein where X is an integer from 1 to 8). Sequences of the invention

are useful for treating or preventing a MOLX-associated disorder in

humans. They are useful for treating or preventing cardiomyopathy,

atherosclerosis and disorders related to cell signal processing and

metabolic pathway modulation. The MOLX antibodies are useful for

treating or preventing diabetes and disorders related to cell signal

processing and metabolic pathway modulation. MOLX sequences are useful

for the treatment or diagnosis of other MOLX-associated disorders, e.g.

chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,

scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral

thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune

diseases, immunodeficiency, HIV (human immuno deficiency virus), viral,

bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,

haematopoietic, thrombolytic, haemostatic diseases, thrombocytopaenia,

ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,

ocular disease, muscular diseases, growth disorders, loss of libido,

stress, depression, pain and epilepsy. They are useful for preventing

chemotherapy side effects and as contraceptives. Sequences of the

invention are also useful for gene therapy. The present sequence

is human kappa casein precursor-like protein, MOL6.

Sequence 181 AA;

Query Match 73.9%; Score 225.5; DB 23; Length 181;

Best Local Similarity 81.8%; Pred. No. 1.1e-19;

Matches 45; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

RESULT 4
AR92150

ID XX AAR92150 standard; Protein; 182 AA.
 XX AC AAR92150;
 XX DT 24-OCT-1996 (first entry)
 XX DE Human milk kappa-casein, inhibits rotaviral infection.
 XX KW Milk protein; kappa-casein; rotavirus inhibition; gastroenteritis;
 KW KW infant formulation; immunodeficiency; diarrhoea.
 XX OS Homo sapiens.
 XX PN WO9608269-A1.
 XX PD 21-MAR-1996.
 XX PF 05-MAY-1995; 95WO-US05676.
 XX 16-SEP-1994; 94US-0308883.
 XX 16-SEP-1994; 94US-0308882.
 XX (ABBO) ABBOTT LAB.
 XX PA Baxter JH, Cummings RD, Mukerji P, Prieto PA, Seo AE;
 XX DR WPI; 1996-179724/18.
 XX KW Kappa-casein used as anti-rotaviral infection agent in nutritional
 PT product - to prevent, retard or treat rotavirus infection,
 PT especially in infants, children and immuno-deficient patients
 XX Claim 1; Fig 9; 42pp; English.
 XX AAR92150 is human milk kappa-casein which is useful as a component
 CC of an enteral nutritional product. Kappa-casein has anti-rotaviral
 CC activity, it inhibits the attachment of human rotavirus to mammalian
 CC cells. The kappa-casein is useful in the treatment and prevention of
 CC onset of gastroenteritis and other diarrhoeal diseases caused by
 CC rotaviridae esp. in infants, children or immunodeficient patients.
 CC Kappa-casein appeared to have no side effects on the subjects treated
 CC and is unlikely to be toxic or cause an allergic reaction.
 XX Sequence 182 AA;
 XX
 Query Match 100.0%; Score 305; DB 17; Length 182;
 Best Local Similarity 100.0%; Pred No. 2.3e-29;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YORRPAIAINNPVPTYYANPVRPQAIPQRYLPNSHPPTVVRRLHPSPF 55
 DB 63 YORRPAIAINNPVPTYYANPVRPQAIPQRYLPNSHPPTVVRRLHPSPF 117
 RESULT 5
 AAE18216
 ID AAE18216 standard; Protein; 181 AA.
 XX AC AAE18216;
 XX DT 07-MAY-2002 (first entry)
 XX DE Human MOL6 protein.
 XX Secreted molecule; MOL6 protein; MOLX; cardiomyopathy; atherosclerosis;
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KW HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;
 KW haematopoietic; thrombolytic; thrombocytopaenia; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;

KW vulnerary; osteopathic; haemostatic; tranquiliser; antidepressant;
 KW analgesic; vasotropic; hypotensive; gene therapy.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /label= Signal_peptide
 FT Protein 25..181
 FT /label= Mature_MOL6_protein
 XX WO200206339-A2.
 XX 24-JAN-2002.
 XX 03-JUL-2001; 2001WO-US21249.
 XX 03-JUL-2000; 2000US-215854P.
 PR 03-JUL-2000; 2000US-215856P.
 PR 03-JUL-2000; 2000US-215902P.
 PR 07-JUL-2000; 2000US-216585P.
 PR 07-JUL-2000; 2000US-216586P.
 PR 17-JUL-2000; 2000US-216722P.
 PR 17-JUL-2000; 2000US-218622P.
 PR 27-JUL-2000; 2000US-218992P.
 PR 27-JUL-2000; 2000US-221285P.
 PR 14-FEB-2001; 2001US-268734P.
 XX (CURA-) CURAGEN CORP.

Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;
 Paturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;
 Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
 Padigaru M;

WPI: 2002-155038/20.
 N-PSDB; AAD28950.

Nucleic acids encoding secreted polypeptides, designated MOLX
 polypeptides, useful for treating a MOLX-associated disorder, e.g.
 cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -

Claim 1; Page 78; 223pp; English.

The patent discloses nucleic acid sequences encoding novel secreted
 molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL
 protein where X is an integer from 1 to 8). Sequences of the invention
 are useful for treating or preventing a MOLX-associated disorder in
 humans. They are useful for treating or preventing cardiomyopathy,
 atherosclerosis and disorders related to cell signal processing and
 metabolic pathway modulation. The MOLX antibodies are useful for
 treating or preventing diabetes and disorders related to cell signal
 processing and metabolic pathway modulation. MOLX sequences are useful
 for the treatment or diagnosis of other MOLX-associated disorders, e.g.
 chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
 scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
 thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
 diseases, immunodeficiency, HIV (human immuno deficiency virus), viral
 bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
 haematopoietic, thrombolytic, haemostatic diseases, thrombocytopaenia,
 ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
 ocular disease, muscular diseases, growth disorders, loss of libido,
 stress, depression, pain and epilepsy. They are useful for preventing
 chemotherapy side effects and as contraceptives. Sequences of the
 invention are also useful for gene therapy. The present sequence
 is human kappa caesin precursor-like protein, MOL6.

Sequence 181 AA;

Query Match 73.9%; Score 225.5; DB 23; Length 181;
 Best Local Similarity 81.8%; Pred. No. 1.1e-19;
 Matches 45; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 11, 2003, 08:16:37 ; Search time 40.0485 Seconds
(without alignments)
217.984 Million cell updates/sec

Title: US-09-508-095-17

Perfect score: 305

Sequence: 1 YQRPAINNPVPTYYA.....YLPNSHPPTVVRNLPSPF 55

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	305	100.0	55	20 AAW93881	Bifidobacterium bi
2	305	100.0	182	14 AAR39351	Recombinant human
3	305	100.0	182	16 AAR2696	Human kappa-casein
4	305	100.0	182	17 AAR92150	Human milk kappa-c
5	225.5	73.9	181	23 AAE18216	Human Molk protein
6	71.5	23.4	193	22 ABB31899	Peptide #4550 enco
7	71.5	23.4	193	22 ABB37137	Peptide #4643 enco
8	71.5	23.4	193	22 ABB22450	Protein #4449 enco
9	71.5	23.4	193	22 AAM57857	Human brain expres

10	71.5	23.4	193	22 AAW70275	Human bone marrow
11	71.5	23.4	193	22 AAM18102	Peptide #4536 enco
12	71.5	23.4	193	22 AAM30614	Peptide #4651 enco
13	71.5	23.4	193	22 AAM05737	Peptide #4419 enco
14	71.5	23.4	193	23 AAG39918	Human peptide enco
15	71.5	23.4	193	22 AAG06694	Novel human diagno
16	71.5	23.4	904	22 ABL15681	Novel human diagno
17	71.5	23.4	904	22 ABL15681	Novel human diagno
18	70	23.0	405	23 ABL15681	Novel human diagno
19	68	22.3	334	9 AAP82971	Human NOV93a prote
20	68	22.3	334	9 AAP82971	Bloodhesive precu
21	67.5	22.1	99	22 AAO08818	Sequence of a bloa
22	67.5	22.1	759	22 AAB70046	Human polypeptide
23	67.5	22.1	914	18 AAW24800	Drosophila melanog
24	67.5	22.1	1312	19 AAW33807	Spinocerebellar at
25	67.5	22.1	1312	20 AAV33495	Human ataxin-2
26	67.5	22.1	1312	24 ABB82698	Human SCA2 protei
27	67.5	22.1	1313	19 AAW60213	Spinocerebellar at
28	67	22.0	914	22 ABB69998	Drosophila melanog
29	66.5	21.8	273	20 AAY50557	HSV1 LAT protein.
30	66.5	21.8	556	22 AAG05032	Novel human diagno
31	66	21.6	580	22 AAB60493	Human cell cycle a
32	65	21.3	112	21 AAY84457	Amino acid sequenc
33	64	21.0	82	22 AAP61243	Propionibacterium
34	64	21.0	86	9 AAP82973	Bloodhesive precu
35	64	21.0	86	9 AAP82973	Sequence of a bloa
36	64	21.0	707	22 ABB68526	Drosophila melanog
37	64	21.0	751	16 AAR80839	Japanese sea musse
38	64	21.0	899	22 ABB25235	Novel human diagno
39	64	21.0	1729	22 ABB62538	Drosophila melanog
40	63.5	20.8	351	23 AAG70805	Human myocardin 3
41	63.5	20.8	351	24 AAG74177	Human cardiac-spec
42	63	20.7	203	9 AAP82970	Bloodhesive precu
43	63	20.7	203	9 AAP82970	Sequence of a bloa
44	63	20.7	511	22 ABB64793	Drosophila melanog
45	63	20.7	652	9 AAP82974	Bloodhesive precu

ALIGNMENTS

RESULT 1

AAW93881

ID AAW93881 standard; peptide; 55 AA.

AC AAW93881;

XX

DT 25-JUN-1999 (first entry)

XX

DE Bifidobacterium bifidus stimulating peptide 17.

XX

KW Bifidogenic peptide; protease; treatment; microbe-related disease;
bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;
infection; inflammation; microbial induced tumour; degenerative disorder;
diarrhoea; colic; oral microflora; intestinal microflora; caries;
vaginal microflora.

OS Bifidobacterium bifidus.

XX

PN WO9914231-A2.

XX

PD 25-MAR-1999.

XX

PF 16-SEP-1998; 98WO-EP05899.

XX

PR 11-FEB-1998; 98DE-1005385.

XX

PR 16-SEP-1997; 97DE-1040604.

XX

FA (FORS/) FORSMANN W.

XX

PI Forssmann W, Liepke C, Zucht H;

XX

DR WPI; 1999-244022/20.

Mon Aug 11 08:31:57 2003

XX Milk-derived peptides that stimulate Bifidobacterium bifidus

PT Claim 2; Page 3; 25pp; German.

PS

XX This invention describes milk-derived bifidogenic peptides and their

CC active derivatives or fragments, and combinations of them produced by

CC chemical coupling. Such are produced from bovine or human milk by

CC treatment for 2 hr with proteases, then centrifuging to remove fat and

CC acidifying to pH 2 to precipitate proteins. The solution phase is then

CC subjected to reverse-phase high-performance liquid chromatography (HPLC)

CC and cation-exchange HPLC, the fractions adjusted to salt content below

CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by

CC growing Bifidobacterium bifidus and Escherichia coli in presence of the

CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15

CC are selected where Bw = germ count after 16 hr culture of B. bifidus in

CC 50% Elliker broth containing peptide at 0.2 mg/ml, B0 = germ count after 16

CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at

CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free

CC control. The peptides AAW93865-W93888 are used to treat microbe-related

CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,

CC mycoplasma, filaria and plasmidia, e.g. infections, inflammation, colic,

CC microbially induced tumours or degenerative disorders, diarrhoea, colic,

CC abnormalities in oral, intestinal or vaginal microflora, or caries.

XX Sequence 55 AA;

Query Match 100.0%; Score 305; DB 20; Length 55;

Best Local Similarity 100.0%; Pred. No. 5.7e-30;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 55

DB 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 55

RESULT 2

AAR39351

ID AAR39351 standard; Protein; 182 AA.

XX AAR39351;

XX 25-MAR-2003 (updated)

DT 26-JAN-1994 (first entry)

DE Recombinant human kappa casein.

XX Casein; supplement; milk; pharmaceutical; ss.

XX Homo sapiens.

XX WO9315196-A1.

XX 05-AUG-1993.

XX 25-JAN-1993; 93WO-DK00024.

XX 23-JAN-1992; 92DK-0000088.

XX (SYMB-) SYMBICOM AB.

XX Bergstroem S, Hansson L, Hernell O, Stroemqvist M;

PI Toernell J;

XX WPI; 1993-258675/32.

DR N-PSDB; AAQ46850.

XX DNA encoding human kappa-casein - used for obtaining recombinant

PT polypeptide(s) for use as nutrient supplements, partic. in infant

PT formulae

XX Claim 44; Page 88-89; 110pp; English.

PS

XX The recombinant human kappa casein is produced in high yields by

CC means of either a eukaryotic or prokaryotic expression system. It

CC is used as a nutrient supplement in milk based products to provide a

CC substantial improvement of the nutritional and biological value of

CC the formulae, making it closer in similarity to human milk. It can

CC also be used as a pharmaceutical.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 182 AA;

Query Match 100.0%; Score 305; DB 14; Length 182;

Best Local Similarity 100.0%; Pred. No. 2.3e-29;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 55

DB 63 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 117

RESULT 3

AAR72696

ID AAR72696 standard; Protein; 182 AA.

XX AAR72696;

XX 06-NOV-1995 (first entry)

DT Human kappa-casein.

DE Kappa-casein; milk protein.

XX Homo sapiens.

XX US5391497-A.

XX 21-FEB-1995.

XX 13-OCT-1992; 92US-0962569.

XX 13-OCT-1992; 92US-0962569.

XX (COLS) UNIV COLORADO FOUND INC.

XX Chang Y, Ham RG, Jeffers KF, Menon RS;

XX WPI; 1995-160470/21.

DR N-PSDB; AAQ89598.

XX DNA encoding human kappa-casein - used for the prodn. of large amts. of

PT highly purified kappa-casein milk protein for infant use.

XX Disclosure; Column 13-16; 14pp; English.

XX A commercial cDNA library prep. in lambda gtl1 from mRNA obtd. from

CC human breast tissue removed during the third trimester of pregnancy

CC was screened with rabbit anti-bovine kappa-casein cDNA. The cDNA

CC insert of a recombinant phage was amplified by PCR to obtain a

CC full-length clone (AAQ89598) encoding human kappa-casein (AAR72696).

XX Sequence 182 AA;

Query Match 100.0%; Score 305; DB 16; Length 182;

Best Local Similarity 100.0%; Pred. No. 2.3e-29;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 55

DB 63 YORRPAIANNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVVRPNLHPSF 117

RESULT 4

AAR92150

ID AAR92150 standard; Protein; 182 AA.
 AC AAR92150;
 XX
 XX
 DT 24-OCT-1996 (first entry)
 XX
 DE Human milk kappa-casein, inhibits rotaviral infection.
 XX
 KW Milk protein; kappa-casein; rotavirus inhibition; gastroenteritis;
 KW infant formulation; immunodeficiency; diarrhoea.
 XX
 OS Homo sapiens.
 XX
 PN WO9608269-A1.
 PD 21-MAR-1996.
 XX
 XX 05-MAY-1995; 95WO-US05676.
 PF
 XX 16-SEP-1994; 94US-0308883.
 PR 16-SEP-1994; 94US-0308882.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Baxter JH, Cummings RD, Mukerji P, Prieto PA, Seo AE;
 XX WPI; 1996-179724/18.
 DR
 XX
 PT Kappa-casein used as anti-rotaviral infection agent in nutritional
 PT product - to prevent, retard or treat rotavirus infection,
 PT especially in infants, children and immuno-deficient patients
 XX
 PS Claim 1; Fig 9; 42pp; English.
 XX
 CC AAR92150 is human milk kappa-casein which is useful as a component
 CC of an enteral nutritional product. Kappa-casein has anti-rotaviral
 CC activity, it inhibits the attachment of human rotavirus to mammalian
 CC cells. The kappa-casein is useful in the treatment and prevention of
 CC onset of gastroenteritis and other diarrhoeal diseases caused by
 CC rotaviridae esp. in infants, children or immunodeficient patients.
 CC Kappa-casein appeared to have no side effects on the subjects treated
 CC and is unlikely to be toxic or cause an allergic reaction.
 XX
 SQ Sequence 182 AA;
 Query Match 100.0%; Score 305; DB 17; Length 182;
 Best Local Similarity 100.0%; Pred. No. 2.3e-29;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YQRRPAIANNPVPRYYANPAVVRPHQIPQOXYLPNSHPPTVVRPNLHPSF 55
 Db 63 YQRRPAIANNPVPRYYANPAVVRPHQIPQOXYLPNSHPPTVVRPNLHPSF 117
 RESULT 5
 AAE18216
 ID AAE18216 standard; Protein; 181 AA.
 AC AAE18216;
 XX
 XX 07-MAY-2002 (first entry)
 DT
 XX
 DE Human MOL6 protein.
 XX
 KW Secreted molecule; MOL6 protein; MOLX; cardiomyopathy; atherosclerosis;
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;
 KW HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;

KW vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;
 KW analgesic; vasotropic; hypotensive; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= Signal_peptide
 FT Protein 25..181
 FT /label= Mature_MOL6_protein
 XX
 PN WO200206339-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 03-JUL-2001; 2001WO-US21249.
 XX
 PR 03-JUL-2000; 2000US-215854P.
 PR 03-JUL-2000; 2000US-215856P.
 PR 03-JUL-2000; 2000US-215902P.
 PR 07-JUL-2000; 2000US-216585P.
 PR 07-JUL-2000; 2000US-216586P.
 PR 07-JUL-2000; 2000US-216722P.
 PR 17-JUL-2000; 2000US-218622P.
 PR 17-JUL-2000; 2000US-218992P.
 PR 27-JUL-2000; 2000US-221285P.
 PR 14-FEB-2001; 2001US-268734P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;
 PI Patturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;
 PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;
 PI Padigaru M;
 XX
 DR WPI; 2002-155038/20.
 DR N-PSDB; AAD28950.
 XX
 PT Nucleic acids encoding secreted polypeptides, designated MOLX
 PT polypeptides, useful for treating a MOLX-associated disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -
 XX
 PS Claim 1; Page 78; 223pp; English.
 XX
 CC The patent discloses nucleic acid sequences encoding novel secreted
 CC molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL
 CC protein where X is an integer from 1 to 8). Sequences of the invention
 CC are useful for treating or preventing a MOLX-associated disorder in
 CC humans. They are useful for treating or preventing cardiomyopathy,
 CC atherosclerosis and disorders related to cell signal processing and
 CC metabolic pathway modulation. The MOLX antibodies are useful for
 CC treating or preventing diabetes and disorders related to cell signal
 CC processing and metabolic pathway modulation. MOLX sequences are useful
 CC for the treatment or diagnosis of other MOLX-associated disorders, e.g.
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune
 CC diseases, immunodeficiency, HIV (human immuno deficiency virus), viral,
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,
 CC ocular disease, muscular diseases, growth disorders, loss of libido,
 CC stress, depression, pain and epilepsy. They are useful for preventing
 CC chemotherapy side effects and as contraceptives. Sequences of the
 CC invention are also useful for gene therapy. The present sequence
 CC is human kappa caesin precursor-like protein, MOL6.
 XX
 SQ Sequence 181 AA;
 Query Match 73.9%; Score 225.5; DB 23; Length 181;
 Best Local Similarity 81.8%; Pred. No. 1.1e-19;
 Matches 45; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Db 15 PAPAHHHP-----EYQGPVSHPHHMPPOQHYPAPPPPPPPPPSHPMFHP 60

RESULT 13

AAW05737
ID AAW05737 standard; Protein; 193 AA.

AC AAW05737;

DT 09-OCT-2001 (first entry)

DE Peptide #4419 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US00661.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -

XX Claim 27; SEQ ID NO 14477; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 193 AA;

XX Query Match

XX Best Local Similarity 23.4%; Score 71.5; DB 22; Length 193;

XX Matches 18; Conservative 4; Mismatches 23; Indels 5; Gaps 2;

QY 5 PATAINNPVPRYYANPAVVRP-HAQIPQRYVLPNSHPTVVRRLNHP 53

Db 15 PAPAHHHP-----EYQGPVSHPHHMPPOQHYPAPPPPPPPSHPMFHP 60

RESULT 14

ABG39918

ID ABG39918 standard; Peptide; 193 AA.

XX ABG39918;

XX

DT

XX

DE

XX

XX

KW

XX

KW

XX

KW

XX

KW

XX

KW

XX

KW

XX

KW

XX

XX

OS

XX

XX

PN

XX

XX

PD

XX

XX

PF

XX

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PA

XX

XX

PI

XX

XX

XX

DR

XX

XX

PT

XX

XX

XX

PS

XX

XX

CC

CC

19-AUG-2002 (first entry)

Human peptide encoded by genome-derived single exon probe SEQ ID 29583.

Human; single exon probe; asthma; lung cancer; COPD; ILD;

chronic obstructive pulmonary disease; interstitial lung disease;

familial idiopathic pulmonary fibrosis; neurofibromatosis;

tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

primary ciliary dyskinesia; pulmonary hypertension;

hyaline membrane disease.

XX

OS

XX

XX

PN

XX

XX

PD

XX

XX

PF

XX

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PA

XX

XX

PI

XX

XX

XX

DR

XX

XX

PT

XX

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	305	100.0	182	1	US-07-962-569A-8	Sequence 8, Appli
2	305	100.0	182	1	US-08-308-883-2	Sequence 2, Appli
3	305	100.0	182	1	US-08-730-163-2	Sequence 2, Appli
4	305	100.0	182	3	US-08-256-799-2	Sequence 2, Appli
5	305	100.0	182	3	US-08-462-437-2	Sequence 2, Appli
6	298	97.7	182	3	US-08-462-437-31	Sequence 31, Appli
7	68	22.3	331	6	5202236-37	Patent No. 5202236
8	68	22.3	334	6	5202236-3	Patent No. 5202236
9	67.5	22.1	1312	3	US-09-041-886-19	Sequence 19, Appli
10	67.5	22.1	1312	4	US-09-648-281-2	Sequence 2, Appli
11	66	21.6	297	4	US-09-252-991A-18170	Sequence 18170, A
12	65.5	21.5	822	4	US-09-252-991A-21920	Sequence 21920, A
13	64	21.0	86	6	5202236-7	Patent No. 5202236
14	64	21.0	691	4	US-09-252-991A-16809	Sequence 16809, A
15	63	20.7	202	6	5202236-36	Patent No. 5202236
16	63	20.7	203	6	5202236-1	Patent No. 5202236
17	63	20.7	652	6	5202236-13	Patent No. 5202236
18	63	20.7	744	6	5202236-25	Patent No. 5202236
19	62.5	20.5	417	4	US-09-252-991A-22460	Sequence 22460, A
20	62.5	20.5	628	4	US-09-252-991A-18780	Sequence 18780, A
21	61.5	20.2	161	4	US-09-252-991A-30695	Sequence 30695, A
22	61	20.0	865	4	US-09-281-766-19	Sequence 19, Appli
23	60.5	19.8	144	4	US-09-252-991A-21500	Sequence 21500, A
24	60.5	19.8	292	4	US-09-252-991A-22366	Sequence 22366, A
25	60.5	19.8	636	4	US-09-252-991A-24902	Sequence 24902, A
26	59.5	19.5	728	4	US-09-252-991A-31845	Sequence 31845, A
27	59.5	19.5	726	4	US-09-252-991A-20875	Sequence 20875, A

```
RESULT 2
US-08-308-883-2
: Sequence 2, Application US/08308883
: Patent No. 5576300
: GENERAL INFORMATION:
: APPLICANT: Mukerji, P. A.
: APPLICANT: Prieto, P. A.
: APPLICANT: Seo, A. E.-Y.
: APPLICANT: Baxter, J. H.
: APPLICANT: Cummings, R. D.
: TITLE OF INVENTION: Method for Inhibition of Human Rotavirus Infection.
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lonnie R. Drayer
: ADDRESSEE: ROSS Products Division
: ADDRESSEE: Abbott Laboratories
: STREET: 625 Cleveland Avenue
: CITY: Columbus
: STATE: Ohio
: COUNTRY: United States
: ZIP: 43215
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb storage
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Macintosh System 7.1
: SOFTWARE: Clarisworks 1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/308,883
: FILING DATE: 16-SEP-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA: No. 5576300 applicable
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (614) 624-3774
: TELEFAX: (614) 624-3074
: TELEX: No. 5576300e
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 182 amino acids
: TYPE: Amino acid
: STRANDEDNESS:
: TOPOLOGY: Linear
: MOLECULE TYPE: Protein.
: DESCRIPTION:
: HYPOTHETICAL:
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
: ORGANISM:
: STRAIN:
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
: ORGANELLE:
: IMMEDIATE SOURCE:
: LIBRARY:
: CLONE:
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT:
: MAP POSITION:
: UNITS:
: FEATURE:
: NAME/KEY:
: LOCATION:
: IDENTIFICATION METHOD:
: OTHER INFORMATION:
: PUBLICATION INFORMATION:
: AUTHORS: L. Hansson et al
: TITLE: DNA Encoding Kappa-Casein, Process for Obtaining the Protein and Use The

: VOLUME:
: ISSUE:
: PAGES:
: DATE:
: DOCUMENT NUMBER: PCT/WO93/15196
: FILING DATE: 25-JAN-1993
: PUBLICATION DATE: 05-AUG-1993
: RELEVANT RESIDUES IN SEQ ID NO:
US-08-308-883-2
Query Match 100.0%; Score 305; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.6e-30;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YORRPAITANNPYVPTVYANPAVVRPQAQIPQRYLPNSHPPTVVRPNLHPSPF 55
DB 63 YORRPAITANNPYVPTVYANPAVVRPQAQIPQRYLPNSHPPTVVRPNLHPSPF 117

RESULT 3
US-08-730-163-2
: Sequence 2, Application US/08730163
: Patent No. 5712250
: GENERAL INFORMATION:
: APPLICANT: Mukerji, P. A.
: APPLICANT: Prieto, P. A.
: APPLICANT: Seo, A. E.-Y.
: APPLICANT: Baxter, J. H.
: APPLICANT: Cummings, R. D.
: TITLE OF INVENTION: Product for Inhibition of Human Rotavirus Infection.
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lonnie R. Drayer
: ADDRESSEE: ROSS Products Division
: ADDRESSEE: Abbott Laboratories
: STREET: 625 Cleveland Avenue
: CITY: Columbus
: STATE: Ohio
: COUNTRY: United States
: ZIP: 43215
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb storage (B)COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Macintosh System 7.1(D)SOFTWARE: Clarisworks 1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/730,163
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/308,882
: FILING DATE: 16-SEP-1994
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (614) 624-3774
: TELEFAX: (614) 624-3074
: TELEX: No. 5712250e
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 182 amino acids
: TYPE: Amino acid
: STRANDEDNESS:
: TOPOLOGY: Linear
: MOLECULE TYPE: Protein.
: DESCRIPTION:
: HYPOTHETICAL:
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
: ORGANISM:
: STRAIN:
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
```

CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: L. Hansson et al
TITLE: DNA Encoding Kappa-Casein, Process for Obtaining the Protein and Use Thereof
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER: PCT/WO93/15196
FILING DATE: 25-JAN-1993
PUBLICATION DATE: 05-AUG-1993
RELEVANT RESIDUES IN SEQ ID NO:
US-08-730-163-2

Query Match 100.0%; Score 305; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.6e-30;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPRYYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 55
|||||
DB 63 YORRPAIANNPVPRYYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 117

RESULT 4
US-08-256-799-2
; Sequence 2, Application US/08256799
; Patent No. 6222094
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: STROEMQVIST, Mats
; APPLICANT: BERGSTROEM, Sven
; APPLICANT: HERNELL, Olle
; APPLICANT: Toernell, Jan
; TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS FOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,799
; FILING DATE: 06-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 88/92
; FILING DATE: 23-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: HANSSON-1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-799-2

Query Match 100.0%; Score 305; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.6e-30;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPRYYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 55
|||||
DB 63 YORRPAIANNPVPRYYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 117

RESULT 5
US-08-462-437-2
; Sequence 2, Application US/08462437
; Patent No. 6232094
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: STROEMQVIST, Mats
; APPLICANT: BERGSTROEM, Sven
; APPLICANT: HERNELL, Olle
; APPLICANT: Toernell, Jan
; TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,437
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 88/92
; FILING DATE: 23-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: HANSSON-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-462-437-2

Query Match 100.0%; Score 305; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.6e-30;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPRYYANPAVVRPHQAIPQRYQLPNSHPPTVVRPNLHPSF 55
|||||

Db 63 YORRPAIANNPVPRYYANPAVRPHQAIPQROYLPNSHPPTVVRRLNHPSF 117

RESULT 6

US-08-462-437-31

Sequence 31, Application US/08462437

Patent No. 6232094

GENERAL INFORMATION:

APPLICANT: HANSSON, Lennart

APPLICANT: STROEMVIST, Mats

APPLICANT: BERGSTROM, Sven

APPLICANT: HERNELL, Olle

APPLICANT: TOERNELL, Jan

TITLE OF INVENTION: DNA ENCODING KAPPA-CASEIN, PROCESS

TITLE OF INVENTION: FOR OBTAINING THE PROTEIN AND USE THEREOF

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,437

FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DK 88/92

FILING DATE: 23-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: COOPER, Iver P.

REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: HANSSON-1A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 182 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-462-437-31

Query Match 97.7%; Score 298; DB 3; Length 182;

Best Local Similarity 98.2%; Pred. No. 3.3e-29;

Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPRYYANPAVRPHQAIPQROYLPNSHPPTVVRRLNHPSF 55

Db 63 YORRPAIANNPVPRYYANPAVRPHQAIPQROYLPNSHPPTVVRRLNHPSF 117

RESULT 7

US-08-462-437-31

Patent No. 5202236

APPLICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG, SUSAN L.; MCCANDLISS, RUSS; WEI, TENA; FILPULA, DAVID

TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE

PROTEIN

NUMBER OF SEQUENCES: 39

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/528,762

FILING DATE: 25-MAY-1990

APPLICATION NUMBER: 82,456

FILING DATE: 07-AUG-1987

APPLICATION NUMBER: 933,945

FILING DATE: 13-SEP-1984

SEQ ID NO: 3

LENGTH: 334

5202236-3

Query Match 22.3%; Score 68; DB 6; Length 334;

Best Local Similarity 29.3%; Pred. No. 0.96;

Matches 17; Conservative 9; Mismatches 28; Indels 4; Gaps 2;

QY 1 YORRPAIANNPVPRYYA---NPVVVRPHQAIPQROYLPNSHPPTVVRRLNHPSF 55

Db 127 YKAKPTYKI-KPTVPSTYKAKPTNPSTYKAKPSYPPTYKAKPSYPPTY 183

RESULT 8

5202236-3

Patent No. 5202236

APPLICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG, SUSAN L.; MCCANDLISS, RUSS; WEI, TENA; FILPULA, DAVID

TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE

PROTEIN

NUMBER OF SEQUENCES: 39

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/528,762

FILING DATE: 25-MAY-1990

APPLICATION NUMBER: 82,456

FILING DATE: 07-AUG-1987

APPLICATION NUMBER: 933,945

FILING DATE: 24-NOV-1986

APPLICATION NUMBER: 650,128

FILING DATE: 13-SEP-1984

SEQ ID NO: 3

LENGTH: 334

5202236-3

Query Match 22.3%; Score 68; DB 6; Length 334;

Best Local Similarity 29.3%; Pred. No. 0.96;

Matches 17; Conservative 9; Mismatches 28; Indels 4; Gaps 2;

QY 1 YORRPAIANNPVPRYYA---NPVVVRPHQAIPQROYLPNSHPPTVVRRLNHPSF 55

Db 130 YKAKPTYKI-KPTVPSTYKAKPTNPSTYKAKPSYPPTYKAKPSYPPTY 186

RESULT 9

US-09-041-886-19

Sequence 19, Application US/09041886

Patent No. 6235872

GENERAL INFORMATION:

APPLICANT: Bredesen, Dale E.

APPLICANT: Rabizadeh, Sharroz

TITLE OF INVENTION: Proapoptotic Peptides, Dependence

TITLE OF INVENTION: Polypeptides and Methods of Use

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell & Flores LLP

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041,886

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:


```

RESULT 11
US-09-252-991A-18170
; Sequence 18170, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18170
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18170

```

RESULT 13
5202236-7
; Patent No. 5202236
; APPLICANT: MAUCH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG,
; SUSAN L.; MCCANDLISS, RUSS; WEI, TENA; FILIPULA, DAVID
; TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE

```

Query Match      21.0%; Score 64; DB 6; Length 86;
Best Local Similarity 28.6%; Pred. NO. 0.61;
Matches 18; Conservative 10; Mismatches 25; Indels 10; Gaps 3;

1 YQRPATAIN---NPXVPTYTYANPAVPAHAQIPQRQYLPN-----SHPTTVVRRRLNH 52
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 YKAKPSYPPTYKAKPTPTSYTKAKPSY---PPYTKPKISYPTYTKAKPSYPTTKAKPSY 61

```

QY 53 PSF 55
Db 62 PTY 64

Search completed: August 11, 2003, 08:19:44
Job time : 22.8932 secs

RESULT 14
US-09-252-991A-16809
; Sequence 16809, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16809
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16809

Query Match 21.0%; Score 64; DB 4; Length 691;
Best Local Similarity 30.9%; Pred. No. 6.9;
Matches 21; Conservative 8; Mismatches 21; Indels 18; Gaps 3;
QY 2 QRRPAIAINNPY-VPRTYTYANPAVVRP-----HAQI-----PQROYLPNSHPP 43
Db 3 RRPAPSPAPPEVPPTTYRHDHRAATPRRAHSGVNCDEPRDLVPSHRRRPLPCHHFG 62
QY 44 TVVVRNPL 51
Db 63 RPCRRPRL 70

RESULT 15
5202236-36
; Patent No. 5202236
; APPLICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG,
; SUSAN L.; MCCANDLISS, RUSS WEL; TENA, FILPULA, DAVID
; TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
; PROTEIN
; NUMBER OF SEQUENCES: 39
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,762
; FILING DATE: 25-MAY-1990
; APPLICATION NUMBER: 82,456
; FILING DATE: 07-AUG-1987
; APPLICATION NUMBER: 933,945
; FILING DATE: 24-NOV-1986
; APPLICATION NUMBER: 650,128
; FILING DATE: 13-SEP-1984
; SEQ ID NO:36;
; LENGTH: 202
5202236-36

Query Match 20.7%; Score 63; DB 6; Length 202;
Best Local Similarity 23.1%; Pred. No. 2.2;
Matches 18; Conservative 12; Mismatches 24; Indels 24; Gaps 3;
QY 1 YORRPAIAINNPYVPRTYTYANPAV-----VRPHAQIPQROYLPN--- 39
Db 70 YKAKPTYKA-KPTYPTTYRAKPSYPTTYKPKSPYPTTKSKSYSPSKKTKYPTTKP 128
QY 40 --SHPTVVRPNLHPSF 55
Db 129 KLTYPPTYKPKSPPSY 146

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:17:06 ; Search time 151.65 Seconds
(without alignments)
46.167 Million cell updates/sec

Title: US-09-508-095-17

Perfect score: 305

Sequence: 1 YQRRPAIAINNPVPTTYA.....YLPNSHPPTVVRPNLHPSPF 55

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 479057 seqs, 127295195 residues

Total number of hits satisfying chosen parameters: 479057

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.5	25.1	413	14	US-10-062-254-270
2	71.5	23.4	193	9	US-09-864-761-37748
3	65	21.3	425	14	US-10-062-254-274
4	63.5	20.8	351	14	US-10-029-217A-7
5	63	20.7	81	15	US-10-083-357-1283
6	61	20.0	865	10	US-09-957-995A-19
7	60.5	19.8	60	10	US-09-816-669A-10
8	58.5	19.2	462	10	US-09-974-298-151
9	58.5	19.2	462	15	US-10-043-487-349
10	58	19.0	366	10	US-09-801-368-254
11	57.5	18.9	402	8	US-08-957-425-29
12	57.5	18.9	402	8	US-08-260-675-21
13	57.5	18.9	402	15	US-10-122-026-6
14	57.5	18.9	402	15	US-10-050-050-21
15	57.5	18.9	403	10	US-09-813-398-31

16	57.5	18.9	420	14	US-10-062-254-254
17	57.5	18.9	509	9	US-09-205-658-46
18	57.5	18.9	509	9	US-09-844-353A-46
19	57.5	18.9	635	9	US-09-844-353A-101
20	57.5	18.9	979	9	US-09-903-187A-5
21	57	18.7	180	9	US-09-739-852-12
22	57	18.7	369	9	US-09-906-779-5
23	57	18.7	662	15	US-10-146-473-65
24	57	18.7	671	15	US-10-128-714-8173
25	56.5	18.5	199	11	US-09-988-067B-40
26	56.5	18.5	315	9	US-09-764-853-772
27	56.5	18.5	315	9	US-09-764-898-261
28	56.5	18.5	315	11	US-09-764-881-98
29	56.5	18.5	315	15	US-10-073-865-87
30	56.5	18.5	363	10	US-09-764-864-1396
31	56.5	18.5	409	14	US-10-062-254-276
32	56.5	18.5	474	9	US-09-752-639-148
33	56.5	18.5	474	10	US-09-984-198-148
34	56.5	18.5	935	10	US-09-764-864-960
35	56.5	18.5	2971	15	US-10-146-473-50
36	56.5	18.5	4349	15	US-10-160-758-15
37	56	18.4	796	9	US-09-205-658-40
38	56	18.4	796	9	US-09-844-353A-40
39	56	18.4	858	9	US-09-205-658-41
40	56	18.4	858	9	US-09-844-353A-41
41	56	18.4	892	9	US-09-205-658-42
42	56	18.4	892	9	US-09-844-353A-42
43	55.5	18.2	171	10	US-09-764-877-1567
44	55.5	18.2	190	9	US-09-864-761-36101
45	55.5	18.2	282	12	US-10-017-161-2200

ALIGNMENTS

RESULT 1

US-10-062-254-270

; Sequence 270, Application US/10062254

; Publication No. US2002013882A1.

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Edgar B

; APPLICANT: Cahoon, Rebecca E

; APPLICANT: Falco, Saverio Carl

; APPLICANT: Fang, Yiwen

; APPLICANT: Hantke, Sabine S.

; APPLICANT: Lee, Jian-Ming

; APPLICANT: Li, Zhongsen

; APPLICANT: Miao, Guo-Hua

; APPLICANT: Morgante, Michele

; APPLICANT: Niu, Xiping

; APPLICANT: Odell, Joan

; APPLICANT: Rafalski, Antoni

; APPLICANT: Sakai, Hajime

; APPLICANT: Zheng, Peizhong

; APPLICANT: Zhu, Qun

; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/062,254

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 09/630,346

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/146511

; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: 60/156006

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/156899

; PRIOR FILING DATE: 1999-09-30

; PRIOR APPLICATION NUMBER: 60/157287

; PRIOR FILING DATE: 1999-10-01

; PRIOR APPLICATION NUMBER: 60/169767

; PRIOR FILING DATE: 1999-12-09

; PRIOR APPLICATION NUMBER: 60/171054

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: 60/172958
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/171515
 ; PRIOR FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: 60/173535
 ; PRIOR FILING DATE: 1999-12-29
 ; NUMBER OF SEQ ID NOS: 375
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 270
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; US-10-062-254-270

Query Match 25.1%; Score 76.5; DB 14; Length 413;
 Best Local Similarity 38.6%; Pred. No. 0.68; 17; Indels 3; Gaps 1;
 Matches 17; Conservative 7; Mismatches 17; Indels 3; Gaps 1;
 QY 2 QRRPAIANNPVYPTTYANPAVVRPQAQIPQOYLPSHPPTV 45
 Db 22 RQSPAVARQQWVPMQYFAAAAMVMPHMLPPQHYAP--PPYV 62

RESULT 2
 US-09-864-761-37748

; Sequence 37748, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 37748
 ; LENGTH: 193
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC002467.1
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
 ; OTHER INFORMATION: EST_HUMAN HT: AA831893.1, EVALUE 7.00e-07
 ; US-09-864-761-37748

Query Match 23.4%; Score 71.5; DB 9; Length 193;
 Best Local Similarity 36.0%; Pred. No. 1.1;
 Matches 18; Conservative 4; Mismatches 23; Indels 5; Gaps 2;

QY 5 PATAINNPYVPTTYANPAVVRP-HAQIPQOYLPSHPPTVVRPNLHP 53
 Db 15 PAPAHHHP-----EQGPVVVSHPHHMPQOQYAPPPPPPISSHMPHP 60

RESULT 3

US-10-062-254-274
 ; Sequence 274, Application US/10062254
 ; Publication No. US20020138882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Edgar B
 ; APPLICANT: Cahoon, Rebecca E
 ; APPLICANT: Falco, Saverio Carl
 ; APPLICANT: Fang, Yiwen
 ; APPLICANT: Hantke, Sabine S.
 ; APPLICANT: Lee, Jian-Ming
 ; APPLICANT: Li, Zhongsen
 ; APPLICANT: Miao, Guo-Hua
 ; APPLICANT: Morgante, Michele
 ; APPLICANT: Niu, Xiping
 ; APPLICANT: Odell, Joan
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Zheng, Peizhong
 ; APPLICANT: Zhu, Qun
 ; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/062,254
 ; CURRENT FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: 09/630,346
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 60/146511
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: 60/156006
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 60/156899
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: 60/157287
 ; PRIOR FILING DATE: 1999-10-01
 ; PRIOR APPLICATION NUMBER: 60/169767
 ; PRIOR FILING DATE: 1999-12-09
 ; PRIOR APPLICATION NUMBER: 60/171054
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: 60/172958
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/171515
 ; PRIOR FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: 60/173535
 ; PRIOR FILING DATE: 1999-12-29

Query Match	20.7%	Score 63	DB 15	Length 81
Best Local Similarity	31.8%	Pred. No. 3.9		
Matches 14	Conservative 7	Mismatches 11	Indels 12	Gaps 2
QY	11	NPXYPTYYANPAVVRHQAIPQRYLPNS	-----	HPPTVY 46
Db	4	SPNLPSHVTPGLVTLPHSHIP	-----	LPTIHLSTVYHHPPDII 43

```

; Sequence 10, Application US/09816669A
; Patent No. US20020137019A1
; GENERAL INFORMATION:
; APPLICANT: GARABEDIAN, Michael

```

```

; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
;
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
;
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
;
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 349
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Shigella Flexneri
; US-10-043-487-349

```

Db 259 KFAIRISNCCITHRASPNSEDEKPEVPPRPVPIPPPPVAPDYRRNSAEVTSYSTSEDEPPK 318
QY 45 VVRRPNLHPS 54
Db 319 VPPREPLSPS 328

```

RESULT 10
US-09-801-368-254
; Sequence 254, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109772.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 254
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-254

```

```

; PRIOR APPLICATION NUMBER: US 60/160,587
;
; PRIOR FILING DATE: 1999-10-20
;
; NUMBER OF SEQ ID NOS: 440
;
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 254
;
; LENGTH: 366
;
; TYPE: prt
;
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-254

```

QY	9	INNPPVPR-----TYANP-----AVRPHAQ 30
Db	44	INLPFVORETNPAYANVAQLATSPQAQSGYCRYAVPPPTYPQQPSYQQAULPYAT 103
QY	31	IPQRQYLFSNHPPTVVRRP-----NLHP 53
Db	104	IPNSGNFPSSPFVNAMVPPEVQFGSFLNTLHP 136

RESULT 11
US-08-957-425-29
; Sequence 29, Application US/08957425
; Publication No. US20030069401A1
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; OZKAYNAK, ENGIN
; KUBERASAMPATH, THANGAVEL
; RUEGER, DAVID C.
; PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,425
; FILING DATE: 24-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,570
; FILING DATE: 21-FEB-1992
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-08-957-425-29
Query Match 18.9%; Score 57.5; DB 8; Length 402;
Best Local Similarity 29.3%; Pred. No. 92;
Matches 22; Conservative 8; Mismatches 22; Indels 23; Gaps 4;
QY 2 QRRPAIANNPVVPTYYANPAVVR-PHA-----QIFQRYVLP-----NS 40
Db 238 QRAP--RSQQPEVVTFFFRASPSPIRTPRAVRLRRRQPKSKNELPQANRLPGIFDDVHGS 295
QY 41 HPPTVVRPNLHPSF 55
Db 296 HGRQVCRRHLYVSF 310
RESULT 12
US-08-260-675-21
; Sequence 21, Application US/08260675
; Publication No. US20030104993A1
; GENERAL INFORMATION:
; APPLICANT: RUEGER, DAVID C
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: PANG, ROY HL
; APPLICANT: COHEN, CHARLES M
; TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA HURWITZ & THIBEAULT
; STREET: 55 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/260,675
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,100
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/922,813
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,274
; FILING DATE: 11-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,764
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESO, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-08-260-675-21

Query Match      18.9%; Score 57.5; DB 8; Length 402;
Best Local Similarity 29.3%; Pred. No. 92;
Matches 22; Conservative 8; Mismatches 22; Indels 23; Gaps 4;

QY 2 QRRPAIANNPYVPRYYANPAVVR-PHA-----OIPORQYLP-----NS 40
Db 238 QRAP--RSQQPFVVTFFRASPSPIRTPRAVRPLRRRQPKKSNELPQANRLPGIFDDVHGS 295
QY 41 HPPTVVRRPNLHPSF 55
Db 296 HGRQVCRHLYVSF 310

RESULT 13
US-10-122-026-6
; Sequence 6, Application US/10122026
; Publication No. US20030105004A1
; GENERAL INFORMATION:
; APPLICANT: JONES, WILLIAM K
; TUCKER, RONALD F
; RUEGER, DAVID C
; OPPERMANN, HERMANN
; OZKAYNAK, ENGIN
; KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
; OF MATTER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Administrator, Testa, Hurwitz &
; Thibault, LLP
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/122,026
; FILING DATE: 29-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,542
; FILING DATE: 13-MAR-1995
; APPLICATION NUMBER: US 08/840,510
; FILING DATE: 31-MAR-1993
; APPLICATION NUMBER: US 08/029,335
; FILING DATE: 04-MAR-1993
; APPLICATION NUMBER: US 07/971,091
; FILING DATE: 03-NOV-1992
; APPLICATION NUMBER: US 07/946,235
; FILING DATE: 16-SEP-1992
; APPLICATION NUMBER: US 07/938,336
; FILING DATE: 28-AUG-1992
; APPLICATION NUMBER: US 07/923,780
; FILING DATE: 31-JUL-1992
; APPLICATION NUMBER: US 07/752,857
; FILING DATE: 30-AUG-1991
; APPLICATION NUMBER: US 07/752,764
; FILING DATE: 30-AUG-1991
; APPLICATION NUMBER: US 07/667,274
; FILING DATE: 11-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMACHO, JENNIFER A.
; REGISTRATION NUMBER: 43,526
; REFERENCE/DOCKET NUMBER: STK-060CN
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-122-026-6

Query Match      18.9%; Score 57.5; DB 15; Length 402;
Best Local Similarity 29.3%; Pred. No. 92;
Matches 22; Conservative 8; Mismatches 22; Indels 23; Gaps 4;

QY 2 QRRPAIANNPYVPRYYANPAVVR-PHA-----OIPORQYLP-----NS 40
Db 238 QRAP--RSQQPFVVTFFRASPSPIRTPRAVRPLRRRQPKKSNELPQANRLPGIFDDVHGS 295
QY 41 HPPTVVRRPNLHPSF 55
Db 296 HGRQVCRHLYVSF 310

RESULT 14
US-10-050-050-21
; Sequence 21, Application US/10050050
; Publication No. US20030125230A1
; GENERAL INFORMATION:
; APPLICANT: COHEN, CHARLES M.
; CHARETTE, MARC F.
; KUBERASAMPATH, THANGAVEL
; RUEGER, DAVID C.
; OPPERMANN, HERMANN
; PANG, ROY H.L.
; OZKAYNAK, ENGIN
; SMART, JOHN E.
; TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING
; PROLIFERATION OF EPITHELIAL CELLS.
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/050,050
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,113
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/445,882
; FILING DATE: 22-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-074DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-050-050-21
```


Query Match 18.9%; Score 57.5; DB 15; Length 402;
Best Local Similarity 29.3%; Pred. No. 92;
Matches 22; Conservative 8; Mismatches 22; Indels 23; Gaps 4;
QY 2 QRRPAIAINNPVPTYYANPAVVR-PHA-----QIPQRYLP-----NS 40
DB 238 GRAP--RSQQPFVVTFFRASPPIRTPAVRPLRRQPKKSNELPQANRLPGIEDDVHGS 295
QY 41 HPPTVVRPNLHPSF 55
DB 296 HGRQVCRHLYVSF 310

RESULT 15
US-09-813-398-31
; Sequence 31, Application US/09813398
; Patent No. US20020169292A1
; GENERAL INFORMATION:
; APPLICANT: Bruce D. Weintraub
; APPLICANT: Mariusz W. Szkudlinski
; APPLICANT: University of Maryland
; TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS
; FILE REFERENCE: UOFMD.003C1
; CURRENT APPLICATION NUMBER: US/09/813,398
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/05908
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/19772
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 403
; TYPE: PRT
; ORGANISM: HOMO SAPIEN
US-09-813-398-31

Query Match 18.9%; Score 57.5; DB 10; Length 403;
Best Local Similarity 29.3%; Pred. No. 92;
Matches 22; Conservative 8; Mismatches 22; Indels 23; Gaps 4;
QY 2 QRRPAIAINNPVPTYYANPAVVR-PHA-----QIPQRYLP-----NS 40
DB 239 GRAP--RSQQPFVVTFFRASPPIRTPAVRPLRRQPKKSNELPQANRLPGIEDDVHGS 296
QY 41 HPPTVVRPNLHPSF 55
DB 297 HGRQVCRHLYVSF 311

Search completed: August 11, 2003, 08:25:41
Job time : 151.65 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:16:37 ; Search time 14.9515 Seconds
(without alignments)
353.763 Million cell updates/sec

Title: US-09-508-095-17
Perfect score: 305
Sequence: 1 YQRPATAINPVPRYYA.....YLPNSHPPTVRRPNLHPSF 55

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	305	100.0	182	1 KKHU	casein kappa precu
2	160	52.5	188	2 A48382	kappa-casein - pig
3	129	42.3	171	1 KKG1	kappa-casein - goa
4	129	42.3	192	2 S15313	kappa-casein - goa
5	127	41.6	190	1 KKB08	kappa-casein precu
6	124	40.7	192	1 KKS8A	kappa-casein precu
7	119	39.0	180	2 JC4955	kappa-casein precu
8	93	30.5	234	2 S12092	kappa-casein precu
9	73.5	24.1	332	2 T06482	probable cell wall
10	69.5	22.8	340	1 WMBEL1	latency-related pr
11	68	22.3	381	2 S52385	cell wall protein
12	65	21.3	289	2 AB2067	hypothetical prote
13	65	21.3	875	2 S23760	polyphenolic adhes
14	65	21.3	1400	2 T52359	hypothetical prote
15	64.5	21.1	254	2 T25073	hypothetical prote
16	64.5	21.1	735	2 T40619	neutral trehalase
17	64	21.0	751	2 S68957	adhesive plaque pr
18	62	20.3	173	2 A72450	hypothetical prote
19	62	20.3	346	2 S19129	proline-rich prote
20	61.5	20.2	607	2 AH2891	thiamin biosynthes
21	61.5	20.2	844	2 D97667	hypothetical prote
22	61	20.0	516	2 H71332	hypothetical prote
23	61	20.0	628	2 S19150	thiamin biosynthes
24	60.5	19.8	238	2 T40820	hypothetical prote
25	60.5	19.8	329	2 T10064	proline-rich prote
26	60.5	19.8	506	2 C87704	cytokinin-induced
27	60.5	19.8	753	2 JQ0532	monooxygenase-rela
28	60	19.7	306	2 T52340	OP protein - Kenne
29	60	19.7	636	2 T03439	cell wall-plasma m
					probable ethylene-

ALIGNMENTS

RESULT 1
KKHU

casein kappa precursor [validated] - human

N;Alternate names: kappa-casein

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1988 #sequence_revision 03-Oct-1995 #text_change 08-Dec-2000

C;Accession: JC4957; A56638; A24001; A16604; S22070

R;Edlund, A.; Johansson, T.; Leidvik, B.; Hansson, L.

Gene 174, 65-69, 1996

A;Title: Structure of the human kappa-casein gene.

A;Reference number: JC4957; MUID:97017129; PMID:8863730

A;Accession: JC4957

A;Molecule type: DNA

A;Residues: 1-182 <EDL>

A;Cross-references: GB:U51899; NID:gl245481; PID:AA050772.1; PID:gl245482

R;Björnstrom, S.; Hansson, L.; Hernell, O.; Loennherdal, B.; Nilsson, A.K.; Stroemqvist

DNA Seq. 3, 245-246, 1992

A;Title: Cloning and sequencing of human kappa-casein cDNA.

A;Reference number: A56638; MUID:93208373; PMID:1296818

A;Accession: A56638

A;Molecule type: mRNA

A;Residues: 1-182 <BER>

A;Cross-references: EMBL:X66417; NID:929675; PID:CAA47048.1; PID:g29676

A;Experimental source: lactating mammary gland

A;Note: sequence extracted from NCBI backbone (NCBIP:128264)

R;Brignon, G.; Chtourou, A.; Ribadeau-Dumas, B.

FEBS Lett. 188, 48-54, 1985

A;Title: Preparation and amino acid sequence of human kappa-casein.

A;Reference number: A24001; MUID:85258132; PMID:4018271

A;Accession: A24001

A;Molecule type: protein

A;Residues: 'E', 25-26, 28-123 <BRI>

A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we h

R;Fiat, A.M.; Jolles, J.; Aubert, J.P.; Loucheux-Lefebvre, M.H.; Jolles, P.

Eur. J. Biochem. 111, 333-339, 1980

A;Title: Localisation and importance of the sugar part of human casein.

A;Reference number: A16604; MUID:81114144; PMID:7460900

A;Accession: A16604

A;Molecule type: protein

A;Residues: 118-160, 'T', 163-172, 'P', 174-177, 'PTTS', 182 <FIA>

C;Comment: This protein is only a minor component in human milk, whereas it is a cows

C;Genetics:

A;Gene: GDB:CSN10; cask; CSN3; Kca

A;Cross-references: GDB:5916337; OMIM:601695

A;Map position: 4q21.1-4q21.1

A;Introns: 18/3; 29/3

A;Note: the first intron occurs before the initiator codon

C;Superfamily: kappa-casein

C;Keywords: glycoprotein; mammary gland; milk; pyroglutamic acid

F.1-22/Domain: signal sequence #status predicted <SIG>

F.23-182/Product: kappa-casein #status experimental <MAT>

F:118-182/Product: caseinoglycopeptide #status experimental <CGP>
 F:23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F:117-118/Cleavage site: Phe-Ile (chymosin) #status experimental
 F:133-143-146-151-167-169-176-181/Binding site: carbohydurate (Thr) (covalent) #status experimental
 F:162/Binding site: carbohydurate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 305; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 2.8e-27;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVRRPNLHPSF 55
 |||||
 Db 63 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVRRPNLHPSF 117
 |||||

RESULT 2
 A48382
 kappa-casein - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
 C:Accession: A48382; A12717
 R:Levine, W.B.; Alexander, L.J.; Hoganson, G.E.; Beattie, C.W.
 Anim. Genet. 23, 361-363, 1992
 A:Title: Cloning and sequencing of the porcine kappa-casein cDNA.
 A:Reference number: A48382; MUID:92367959; PMID:1503275
 A:Accession: A48382
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-188 <LEV>
 A:Cross-references: GB:X51977; GB:S42402; NID:9406774; PIDN:CAA36239.1; PID:9406775
 A:Experimental source: mammary gland
 A:Note: sequence extracted from NCBI backbone (NCBIN:110878, NCBI:110880)
 R:Chobert, J.M.; Mercier, J.C.; Bahy, C.; Haze, G.
 FEBS Lett. 72, 173-178, 1976
 A:Title: Structure primaire du caseinomacropéptide des caseines kappa porcine et humaine
 A:Reference number: A12717; MUID:77068846; PMID:1001463
 A:Accession: A12717
 A:Molecule type: protein
 A:Residues: 118-188 <CHO>
 C:Superfamily: kappa-casein

Query Match 52.5%; Score 160; DB 2; Length 188;
 Best Local Similarity 54.5%; Pred. No. 6.4e-11;
 Matches 30; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

OY 1 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNSHPPTVRRPNLHPSF 55
 |||||
 Db 63 YQRRSAVSNRQPIPIPIYARPVVAGPHQAQKQWQDPNPVPTVARRPRPHASF 117
 |||||

RESULT 3
 KKG7
 kappa-casein - goat
 C:Species: Capra aegagrus hircus (domestic goat)
 C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Dec-2000
 C:Accession: A94479; A90670; A03114
 R:Mercier, J.C.; Addeo, F.; Pelissier, J.P.
 unpublished results, cited by Mercier, J.C., Chobert, J.M., and Addeo, F., FEBS Lett. 72
 A:Description: Comparative study of the amino acid sequences of the caseinomacropéptides
 A:Reference number: A94479
 A:Accession: A94479
 A:Molecule type: protein
 A:Residues: 1-171 <ME1>
 A:Note: the amino-terminal residue appears to be pyrrolidone carboxylic acid, but cyclized
 R:Mercier, J.C.; Addeo, F.; Pelissier, J.P.
 Biochimie 58, 1303-1310, 1976
 A:Title: Structure primaire du caseinomacropéptide de la caséine K caprine.
 A:Reference number: A90670; MUID:77112689; PMID:1016651
 A:Accession: A90670
 A:Molecule type: protein
 A:Residues: 106-171 <ME2>
 A:Note: 119-Val was also found
 A:Note: Ser-151 and Ser-168 are phosphorylated

C:Comment: Chymosin hydrolyzes the peptide bond between Phe-105 and Met-106.
 C:Superfamily: kappa-casein
 C:Keywords: mammary gland; milk; phosphoprotein

Query Match 42.3%; Score 129; DB 1; Length 171;
 Best Local Similarity 44.4%; Pred. No. 1.8e-07;
 Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;

OY 1 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNS-----HPTVVRRLNHL 52
 |||||
 Db 43 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNS-----HPTVVRRLNHL 102
 |||||

OY 53 PSF 55
 ||
 Db 103 LSF 105

RESULT 4
 S15513
 kappa-casein - goat
 C:Species: Capra aegagrus hircus (domestic goat)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 13-Aug-1999
 C:Accession: S15513
 R:Coll, A.; Foich, J.M.; Sanchez, A.
 submitted to the EMBL Data Library, July 1991
 A:Reference number: S15513
 A:Accession: S15513
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-192 <COL>
 A:Cross-references: EMBL:X60763; NID:9977; PIDN:CAA43174.1; PID:9978
 C:Superfamily: kappa-casein
 C:Keywords: phosphoprotein

Query Match 42.3%; Score 129; DB 2; Length 192;
 Best Local Similarity 44.4%; Pred. No. 2e-07;
 Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;

OY 1 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNS-----HPTVVRRLNHL 52
 |||||
 Db 64 YQRRPAIAINNPVPTTYANPAVVRPHQAIPQROYLPNS-----HPTVVRRLNHL 123
 |||||

OY 53 PSF 55
 ||
 Db 124 LSF 126

RESULT 5
 KKB0B
 kappa-casein precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 24-Apr-1984 #sequence_revision 23-Mar-1995 #text_change 16-Jun-2000
 C:Accession: S02076; B23071; S06376; JN0364; A91206; A90033; A90920; A91634; A90592;
 R:Alexander, L.J.; Stewart, A.F.; Mackinlay, A.G.; Kapelinskaya, T.V.; Tkach, T.M.; G
 Eur. J. Biochem. 178, 393-401, 1988
 A:Title: Isolation and characterization of the bovine kappa-casein gene.
 A:Reference number: S02076; MUID:89091174; PMID:3208764
 A:Contents: A variant
 A:Accession: S02076
 A:Molecule type: DNA
 A:Residues: 1-190 <ALE>
 A:Cross-references: EMBL:X14907; NID:9177; PIDN:CAA33034.1; PID:91288078
 R:Stewart, A.F.; Willis, I.M.; Mackinlay, A.G.
 Nucleic Acids Res. 12, 3695-3907, 1984
 A:Title: Nucleotide sequences of bovine alpha-s1- and kappa-casein cDNAs.
 A:Reference number: A93517; MUID:84221403; PMID:6328443
 A:Contents: A variant
 A:Accession: B23071
 A:Molecule type: mRNA
 A:Residues: 1-190 <STE>
 A:Cross-references: GB:X00565; NID:9170; PIDN:CAA25231.1; PID:91364187
 R:Gorodetskii, S.I.; Kaledin, A.S.
 Sov. Genet. 23, 398-404, 1987

R: Gorodetskii, S.I.; Kaledin, A.S.
Genetika 23, 398-404, 1987
A: Title: Analysis of nucleotide sequence of bovine kappa-casein cdna.
A: Reference number: I45875
A: Accession: I45875
A: Molecule type: mRNA
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Residues: 1-156, 'I', 158-168, 'A', 170-173, 'T', 175-190 <GOR>
A: Cross-references: GB:M36641; NID:g162810; PIDN:AAA30433.1; PID:g162811
C: Comment: The sequence shown is the A variant.
C: Genetics:
A: Introns: 19/3; 30/3
C: Superfamily: kappa-casein
C: Keywords: glycoprotein; milk; phosphoprotein
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-190/Product: kappa-casein #status predicted <MAT>
F: 126-127/Cleavage site: Phe-Met (chymosin) #status experimental
F: 152/Binding site: carboxylate (Thr) (covalent) #status experimental
F: 170/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 41.6%; Score 127; DB 1; Length 190;
Best Local Similarity 44.4%; Pred. No. 3.4e-07;
Matches 28; Conservative 7; Mismatches 20; Indels 8; Gaps 1;

QY 1 YQRRPAIANNPYVPTYYANPAVVRPHQAIPQROYLPN-----SHPTVVRRPNLH 52
Db 64 YQRPVALINNFPLPYPAKPAVVRSPACILQWVLSNTPVPAKSCQAOPTAWARHPHPH 123
QY 53 PSF 55
Db 124 LSF 126

RESULT 6
KKSHA
Kappa-casein precursor - sheep
C: Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C: Date: 24-Apr-1984 #sequence, revision 30-Sep-1991 #text_change 22-Jun-1999
C: Accession: S14711; A03113; A90597; S08655
R: Furet, J.P.; Mercier, J.C.; Soulier, S.; Gaye, P.; Hue-Delahaie, D.; Vilotte, J.L.
Nucleic Acids Res. 18, 5286, 1990
A: Title: Nucleotide sequence of ovine kappa-casein cdna.
A: Reference number: S14711; MUID: 90384837; PMID: 2402453
A: Accession: S14711
A: Molecule type: mRNA
A: Residues: 1-192 <FUR>
A: Cross-references: EMBL: X51822; NID: g1293; PIDN: CAA36122.1; PID: g1294
R: Jolles, J.; Schoentgen, F.; Hermann, J.; Alais, C.; Jolles, P.
Eur. J. Biochem. 46, 127-132, 1974
A: Title: The sequence of sheep kappa-casein: primary structure of para-kappa-A-casein
A: Reference number: A91221; MUID: 74309256; PMID: 4605338
A: Accession: A03113
A: Molecule type: protein
A: Residues: 22, 'O', 24-27, 'E', 29-129 <JOL>
R: Jolles, J.; Fiat, A.M.; Schoentgen, F.; Alais, C.; Jolles, P.
Biochim. Biophys. Acta 365, 335-343, 1974
A: Title: The amino acid sequence of sheep kappa-A-casein. II. Sequence studies concern
A: Reference number: A90597; MUID: 75036120; PMID: 4429673
A: Accession: A90597
A: Molecule type: protein
A: Residues: 127-192 <JOL2>
C: Superfamily: kappa-casein
C: Keywords: glycoprotein; mammary gland; milk; phosphoprotein
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-192/Product: kappa-casein #status experimental <MAT>

Query Match 40.7%; Score 124; DB 1; Length 192;
Best Local Similarity 42.9%; Pred. No. 7.5e-07;
Matches 27; Conservative 6; Mismatches 22; Indels 8; Gaps 1;

QY 1 YQRRPAIANNPYVPTYYANPAVVRPHQAIPQROYLPN-----SHPTVVRRPNLH 52
Db 64 YQRPVALINNFPLPYPAKPAVVRSPACILQWVLSNTPVPAKSCQAOPTAWARHPHPH 123

A: Title: Analysis of nucleotide sequence of bovine kappa-casein cdna.
A: Reference number: S06376
A: Contents: B2 variant
A: Accession: S06376
A: Molecule type: mRNA
A: Residues: 1-138, 'G', 140-156, 'T', 158-168, 'A', 170-173, 'T', 175-190 <GOR>
R: Gorodetskii, S.I.; Kershulite, D.D.; Korobko, B.G.
Bioorg. Khim. 9, 1693-1694, 1983
A: Title: Primary structure of cdna of Bos taurus kappa-casein macropeptide.
A: Reference number: JN0364; MUID: 85022828; PMID: 6689612
A: Contents: B2 variant
A: Accession: JN0364
A: Molecule type: mRNA
A: Residues: 92-156, 'I', 158-168, 'A', 170-173, 'T', 175-190 <GO2>
A: Cross-references: GB:M38333; NID: g162806; PIDN: AAA30432.1; PID: g162807
A: Note: the authors translated the codon TTG for residue 100 as Pro and GAA for residue
R: Mercier, J.C.; Brignon, G.; Ribadeau-Dumas, B.
Eur. J. Biochem. 35, 222-235, 1973
A: Title: Structure primaire de la caseine kappa bovine. Sequence complete.
A: Reference number: A91206; MUID: 73216467; PMID: 4577852
A: Contents: B variant
A: Accession: A91206
A: Molecule type: protein
A: Residues: 22-101, 'D', 103-127 <MER>
A: Note: the amino-terminal residue appears to be pyrrolidone carboxylic acid, but cyclized
R: Grosclaude, F.; Mahe, M.F.; Mercier, J.C.; Ribadeau-Dumas, B.
Ann. Genet. Sel. Anim. 4, 515-521, 1972
A: Title: Localisation des substitutions d'acides amines differenciant les variants A et
A: Reference number: A90033
A: Contents: A variant
A: Accession: A90033
A: Molecule type: protein
A: Residues: 128-190 <GRO>
A: Note: article in French with English abstract
R: Jolles, J.; Schoentgen, F.; Alais, C.; Jolles, P.
Chimia 26, 645-646, 1972
A: Reference number: A90920
A: Contents: A variant
A: Accession: A90920
A: Molecule type: protein
A: Residues: 22, 'Q', 24-25, 'E', 27, 'E', 29-126 <JO2>
R: Jolles, J.; Schoentgen, F.; Alais, C.; Fiat, A.M.; Jolles, P.
Helv. Chim. Acta 55, 2872-2883, 1972
A: Title: Studies on the primary structure of cow kappa-casein. Structural features of pa
A: Reference number: A91634; MUID: 73124636; PMID: 4653404
A: Contents: B variant
A: Accession: A91634
A: Molecule type: protein
A: Residues: 22-101, 'D', 103-156, 'I', 158-168, 'A', 170-190 <JOL>
R: Guerin, J.; Alais, C.; Jolles, J.; Jolles, P.
Biochim. Biophys. Acta 351, 325-332, 1974
A: Title: kappa-Casein from bovine colostrum.
A: Reference number: A90592; MUID: 74269749; PMID: 4407313
A: Accession: A90592
A: Molecule type: protein
A: Residues: 127-156, 'I', 158-168; 187-190 <GUE>
R: Mercier, J.C.; Uro, J.; Ribadeau-Dumas, B.; Grosclaude, F.
Eur. J. Biochem. 27, 535-547, 1972
A: Title: Structure primaire du caseinomacropéptide de la caseine kappa-1 bovine.
A: Reference number: A60833; MUID: 72254481; PMID: 4559180
A: Contents: B variant
A: Accession: A60833
A: Molecule type: protein
A: Residues: 128-156, 'I', 158-168, 'A', 170-190 <ME2>
A: Note: article in French with English abstract
R: Drohse, H.B.; Foltmann, B.
Biochim. Biophys. Acta 995, 221-224, 1989
A: Title: Specificity of milk-clotting enzymes towards bovine kappa-casein.
A: Reference number: S03993; MUID: 89207571; PMID: 2495817
A: Accession: S03993
A: Molecule type: protein
A: Residues: 126-130 <DRO>

Search completed: August 11, 2003, 08:18:57
Job time : 16.9515 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 11, 2003, 08:16:36 ; Search time 8.00971 Seconds
(without alignments)
322.917 Million cell updates/sec

Title: US-09-508-095-17
Perfect score: 305
Sequence: 1 YRRPATAINPNVPTYYA.....YLPNSHPPTVVRPNLHPSF 55

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	305	100.0	182	1	CASK_HUMAN	P07498	homo sapien
2	187	61.3	135	1	CASK_EOGR	Q28400	equus grevy
3	186	61.0	145	1	CASK_BALPH	Q27952	balaeonopter
4	183	60.0	145	1	CASK_HORSE	P82187	equus caball
5	181	59.3	145	1	CASK_TAPIN	Q29135	caprius ind
6	181	59.3	145	1	CASK_UNCUN	Q29150	uncia uncia
7	171	56.1	145	1	CASK_HIPAM	Q28441	hipopotamu
8	165	54.1	125	1	CASK_LAMGU	Q28451	lama guanac
9	165	54.1	182	1	CASK_CAMDR	P79139	camelus dro
10	160	52.5	146	1	CASK_TAVTA	Q28794	tayassu taj
11	160	52.5	188	1	CASK_PIG	P11841	sus scrofa
12	135	44.3	190	1	CASK_BUBBU	P11840	bubalus bub
13	134	43.9	192	1	CASK_OREAM	P50423	oreamos am
14	129	42.3	192	1	CASK_CAPCR	P42156	capricornis
15	129	42.3	192	1	CASK_CAPHI	P02670	capra hircu
16	129	42.3	192	1	CASK_CAPSU	P50420	capricornis
17	129	42.3	192	1	CASK_CAPSW	P50421	capricornis
18	129	42.3	192	1	CASK_NEMGO	P50422	memoriaedus
19	129	42.3	202	1	CASK_SALTA	P50425	saiga tatar
20	127	41.6	136	1	CASK_BISBO	P42155	bison bonas
21	127	41.6	190	1	CASK_BOVIN	P02668	bos taurus
22	124	40.7	192	1	CASK_RUPRU	P50434	rupicapra r
23	124	40.7	192	1	CASK_SHEEP	P02669	ovis aries
24	121	39.7	190	1	CASK_CERNI	P42157	cervus nipp
25	120	39.3	153	1	CASK_GIRCA	Q28417	giraffa cam
26	119	39.0	180	1	CASK_RABIT	P33618	oryctolagus
27	114	37.4	171	1	CASK_TRAJA	Q29137	tragulus ja
28	111	36.4	122	1	CASK_CERDU	Q29147	cervus duva
29	111	36.4	122	1	CASK_CERUN	Q95177	cervus unia
30	106	34.8	124	1	CASK_OVIMO	Q95227	ovibos mosc
31	103	33.8	122	1	CASK_ELADA	Q95184	elaphurus d
32	103	33.8	122	1	CASK_ODOHE	Q95225	odocoileus
33	103	33.8	122	1	CASK_ODOVI	Q95228	odocoileus

34	103	33.8	122	1	CASK_RANTA	Q95239	rangifer ta
35	101	33.1	124	1	CASK_OVIDA	Q95224	ovis dalli
36	100	32.8	122	1	CASK_MUNRE	Q95199	muntiacus r
37	99	32.5	122	1	CASK_CEREL	Q95149	cervus elap
38	99	32.5	123	1	CASK_CAPCA	Q95146	capreolus c
39	95	31.1	122	1	CASK_WAZAM	Q95191	mazama amer
40	93	30.5	234	1	CASK_CAVPO	P19442	cavia porce
41	72.5	23.8	872	1	FPL1_MYTGO	Q25434	mytilus cor
42	69.5	22.8	340	1	LRP1_HSVIF	P17588	herpes simp
43	69	22.6	728	1	TREB_NEUCR	Q42783	neutrospora
44	65	21.3	875	1	FPL1_MYTED	Q25460	mytilus edu
45	64.5	21.1	735	1	TREA_SCHPO	Q42893	schizosacch

ALIGNMENTS

RESULT 1
CASK_HUMAN
ID CASK_HUMAN STANDARD: PRT: 182 AA.
AC P07498; Q13575;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein precursor.
GN CSN3 OR CSN10 OR CSNK OR CASK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [J]
RP SEQUENCE FROM N.A.
RA Menon R.S., Jeffers K.F., Chang Y.F., Ham R.G.;
RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=93208373; PubMed=1296818;
RA Bergstrom S., Hansson L., Hernell O., Loenneker B., Nilsson A.K.,
RA Stroemqvist M.;
RL "Cloning and sequencing of human kappa-casein cDNA";
RL DNA Seq. 3:245-246(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97017129; PubMed=8863730;
RA Edlund A., Johansson T., Leidvik B., Hansson L.;
RL "Structure of the human kappa-casein gene";
RL Gene 174:65-69(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.L., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Richards S., Morley K.C., McKernan K.J., Hale S., Garcia A.M., Gay L.J.,
RA Villalón D.K., Muzny D.J., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [15]

OS Balaenoptera physalus (Finback whale) (Common rorqual).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
 CC Balaenopteridae; Balaenoptera.
 CC NCBI_TaxID=9770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96364219; PubMed=8752004;
 RA Gatesy J., Hayashi C., Cronin M.A., Arctander P.;
 RT "Evidence from milk casein genes that cetaceans are close relatives
 of hippopotamid artiodactyls.";
 RL Mol. Biol. Evol. 13:954-963(1996).
 CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
 casein precipitation in milk.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -!- SIMILARITY: Belongs to the kappa-casein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U53888; AAB08404.1;
 DR InterPro: IPR000117; Casein_kappa.
 DR Pfam; PF00997; casein_kappa; 1.
 DR ProDom; PD003689; Casein_kappa; 1.
 KW Milk; Phosphorylation; Glycoprotein.
 FT NON_TER 1
 FT SITE .81 82 CLEAVAGE (BY CHYMOSIN (=RENNIN))
 FT SITE (BY SIMILARITY).
 SQ SEQUENCE 145 AA; 16324 MW; 17024F195EA78781 CRC64;
 Query Match 61.0%; Score 186; DB 1; Length 145;
 Best Local Similarity 61.8%; Pred. No. 2e-14;
 Matches 34; Conservative 4; Mismatches 17; Indels 0; Gaps 0;
 QY 1 YORPATAINPVPRYYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSP 55
 DB 27 YQRLAALINQMPYPIYAKPVAVRPHQAIPQWQLPNTHPTVAAHHHPHPSF 81
 RESULT 4
 ID CASK_HORSE STANDARD; PRT; 185 AA.
 AC P82187; Q8SPR0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kappa casein precursor (Kappa-CN).
 GN CSN3 OR CSN10.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 CC NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lenasi T., Rogelj I., Dovc P.;
 RT "Equus caballus kappa-casein (k-CN) mRNA.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 21-181.
 RC TISSUE=Milk;
 RA Tedeschi G.;
 RL Submitted (NOV-1999) to the SWISS-PROT data bank.
 CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
 casein precipitation in milk.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -!- SIMILARITY: Belongs to the kappa-casein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AY040863; AAK83669.1;
 DR InterPro: IPR000117; Casein_kappa.
 DR Pfam; PF00997; casein_kappa; 1.
 DR ProDom; PD003689; Casein_kappa; 1.
 KW Milk; Phosphorylation; Signal.
 FT SIGNAL 1
 FT CHAIN 21 185 KAPPA CASEIN.
 FT SITE 117 118 CLEAVAGE (BY CHYMOSIN (=RENNIN))
 FT SITE (BY SIMILARITY).
 FT CONFLICT 29 34 TCHKND -> RCVKNH (IN REF. 2).
 FT CONFLICT 112 112 C -> R (IN REF. 2).
 FT CONFLICT 152 152 V -> VV (IN REF. 2).
 SQ SEQUENCE 185 AA; 21021 MW; F7634F52CEE90606 CRC64;
 Query Match 60.0%; Score 183; DB 1; Length 185;
 Best Local Similarity 63.6%; Pred. No. 5.7e-14;
 Matches 35; Conservative 4; Mismatches 16; Indels 0; Gaps 0;
 QY 1 YORPATAINPVPRYYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSP 55
 DB 63 YQRLAALINQMPYPIYAKPVAVRPHQAIPQWQLPNTHPTVAAHHHPHPSF 117
 RESULT 5
 ID CASK_TAPIN STANDARD; PRT; 146 AA.
 AC Q29135;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kappa casein (fragment).
 GN CSN3 OR CSN10 OR CSNK.
 OS Tapirus indicus (Asiatic tapir) (Malayan tapir).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.
 CC NCBI_TaxID=9802;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96364219; PubMed=8752004;
 RA Gatesy J., Hayashi C., Cronin M.A., Arctander P.;
 RT "Evidence from milk casein genes that cetaceans are close relatives
 of hippopotamid artiodactyls.";
 RL Mol. Biol. Evol. 13:954-963(1996).
 CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
 casein precipitation in milk.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -!- SIMILARITY: Belongs to the kappa-casein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U53892; AAB08418.1;
 DR InterPro: IPR000117; Casein_kappa.
 DR Pfam; PF00997; casein_kappa; 1.
 DR ProDom; PD003689; Casein_kappa; 1.
 KW Milk; Phosphorylation; Glycoprotein.
 FT NON_TER 1
 FT SITE .80 81 CLEAVAGE (BY CHYMOSIN (=RENNIN))

```

FT SQ SEQUENCE 146 AA; 16691 MW; E8957001E0AD59CF CRC64;
Query Match 59.3%; Score 181; DB 1; Length 146;
Best Local Similarity 61.8%; Pred. No. 7.4e-14;
Matches 34; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQOYLPSHPTVVRPNLHPSF 55
DB 27 YQHPVPVNNQMPYRYARPAVVRPHQAIPQOVLPTNPTVVRHPRPSF 81

RESULT 6
CASK_UNCUN STANDARD; PRT; 146 AA.
AC Q29150;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein (Fragment).
GN CSN3 OR CSN10 OR CSNK.
OS Uncia uncia (Snow Leopard) (Panthera uncia).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Uncia.
OX NCBI_TaxID=29064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96364219; PubMed=8752004;
RA Gatesy J., Hayashi C., Cronin M.A., Arcander P.;
RT "Evidence from milk casein genes that cetaceans are close relatives
of hippopotamid artiodactyls.";
RL Mol. Biol. Evol. 13:954-963(1996).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U53894; AAB08421.1;
CC InterPro: IPR000117; Casein_kappa.
CC Pfam: PF00997; casein_kappa; 1.
CC ProDom: PD003689; Casein_kappa; 1.
CC Milk; Phosphorylation; Glycoprotein.
CC NON_TER 1
CC CARBOHYD 91
CC SEQUENCE 146 AA; 16337 MW; D667DEF607E5B66 CRC64;
Query Match 59.3%; Score 181; DB 1; Length 146;
Best Local Similarity 58.2%; Pred. No. 7.4e-14;
Matches 32; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQOYLPSHPTVVRPNLHPSF 55
DB 27 YPHKPAVNNQMPYRYAKPAVVRPHVQIPQOVLPTNPTVVRHPLPSF 81

RESULT 7
CASK_HIPAM STANDARD; PRT; 145 AA.
AC Q28441;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein (Fragment).
GN CSN3 OR CSN10 OR CSNK.
OS Hippopotamus amphibius (Hippopotamus).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
OX NCBI_TaxID=9833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96364219; PubMed=8752004;
RA Gatesy J., Hayashi C., Cronin M.A., Arcander P.;
RT "Evidence from milk casein genes that cetaceans are close relatives
of hippopotamid artiodactyls.";
RL Mol. Biol. Evol. 13:954-963(1996).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U53899; AAB08413.1;
CC InterPro: IPR000117; Casein_kappa.
CC Pfam: PF00997; casein_kappa; 1.
CC ProDom: PD003689; Casein_kappa; 1.
CC Milk; Phosphorylation; Glycoprotein.
CC NON_TER 1
CC CARBOHYD 91
CC SEQUENCE 145 AA; 16272 MW; 874A9EEFD6F80398 CRC64;
Query Match 56.1%; Score 171; DB 1; Length 145;
Best Local Similarity 60.0%; Pred. No. 1e-12;
Matches 33; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 YORRPAIANNPVPTTYANPAVVRPHQAIPQOYLPSHPTVVRPNLHPSF 55
DB 27 YQCRPAALNNQMPYRYAKPAVVRPHQAIPQOVLPTNPTVVRHPRPSF 81

RESULT 8
CASK_LAMGU STANDARD; PRT; 125 AA.
AC Q28451;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein (Fragment).
GN CSN3 OR CSN10 OR CSNK.
OS Lama guanicoe (Guanaco).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
OX NCBI_TaxID=9840;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96364219; PubMed=8752004;
RA Gatesy J., Hayashi C., Cronin M.A., Arcander P.;
RT "Evidence from milk casein genes that cetaceans are close relatives
of hippopotamid artiodactyls.";
RL Mol. Biol. Evol. 13:954-963(1996).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U53890; AAB08415.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
KW Milk; Phosphorylation; Glycoprotein.
FT NON_TER 1
FT SITE 59 60 CLEAVAGE (BY CHYMOSIN (-RENNIN))
FT (BY SIMILARITY)
SQ SEQUENCE 125 AA; 13795 MW; 9E51E7A07EC8400B CRC64;

Query Match 54.1%; Score 165; DB 1; Length 125;
Best Local Similarity 58.2%; Pred. No. 4e-12;
Matches 32; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

OY 1 YQRRPAIANNPYRTYANPAVVRPHQAIPQOYLPSNHPPTVVRPNLHPSF 55

DB 6 YQHLRAVPIINQFIPYPNYAKPVAIRLHAQIPQCALPNIDPPTVRRPRPSPF 60

RESULT 9

CC CASK_CAMDR
ID CASK_CAMDR STANDARD; PRT; 182 AA.
AC P79139;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein precursor.
DE CSN3 OR CSN10 OR CSNK.
GN Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-Somali; TISSUE-Udder;
RX MEDLINE=98291310; PubMed=9627840;
RA Kappeler S., Farah Z., Puhar Z.;
RT "Sequence analysis of Camelus dromedarius milk caseins.";
RL J. Dairy Res. 65:209-222(1998).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
CC casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Y10082; CAA71171.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
KW Milk; Phosphorylation; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 182 KAPPA CASEIN.
FT SITE 117 118 CLEAVAGE (BY CHYMOSIN (-RENNIN)) (BY
SIMILARITY).
SQ SEQUENCE 182 AA; 20417 MW; 418D19E061DA7338 CRC64;

Query Match 54.1%; Score 165; DB 1; Length 182;
Best Local Similarity 58.2%; Pred. No. 6.1e-12;
Matches 32; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

OY 1 YQRRPAIANNPYRTYANPAVVRPHQAIPQOYLPSNHPPTVVRPNLHPSF 55

DB 63 YQHLRAVPIINQFIPYPNYAKPVAIRLHAQIPQCALPNIDPPTVRRPRPSPF 117

RESULT 10

CC CASK_TAYTA
ID CASK_TAYTA STANDARD; PRT; 146 AA.
AC Q28794;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein (Fragment).
DE CSN3 OR CSN10 OR CSNK.
GN Tayassu tajacu (Collared peccary) (Pecari tajacu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Tayassuidae; Pecari.
OX NCBI_TaxID=9829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96364219; PubMed=8752004;
RA Gatesy J., Hayashi C., Cronin M.A., Arcander P.;
RT "Evidence from milk casein genes that cetaceans are close relatives
RT of hippopotamid artiodactyls.";
RL Mol. Biol. Evol. 13:954-963(1996).
CC -!- FUNCTION: kappa casein stabilizes micelle formation, preventing
CC casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U53891; AAB08416.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
KW Milk; Phosphorylation; Glycoprotein.
FT NON_TER 1
FT SITE 80 81 CLEAVAGE (BY CHYMOSIN (-RENNIN))
FT (BY SIMILARITY).
SQ SEQUENCE 146 AA; 16405 MW; 133188256EC038 CRC64;

Query Match 52.5%; Score 160; DB 1; Length 146;
Best Local Similarity 54.5%; Pred. No. 1.8e-11;
Matches 30; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

OY 1 YQRRPAIANNPYRTYANPAVVRPHQAIPQOYLPSNHPPTVVRPNLHPSF 55

DB 27 YQHSRAASTNNQFIPYPFYARPVAVRPHQAIPQOYLPSNHPPTVVRPNLHPSF 81

RESULT 11

CC CASK_PIG
ID CASK_PIG STANDARD; PRT; 188 AA.
AC P11841;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein precursor.
DE CSN3 OR CSN10 OR CSNK.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Mammary gland;
RX MEDLINE=92367959; PubMed=1503275;


```

CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D32185; BAA06886.1;
DR EMBL; D32183; BAA06886.1; JOINED.
DR EMBL; D32184; BAA06886.1; JOINED.
DR EMBL; U37512; AAC48658.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
DR Milk; Phosphorylation; Glycoprotein; Signal;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 192 KAPPA CASEIN.
FT SITE 126 127 CLEAVAGE (BY CHYMOSIN (-RENNIN))
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 152 152 O-LINKED (GALNAC... ) (BY SIMILARITY).
FT MOD_RES 172 172 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 189 189 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 192 AA; 21525 MW; 95BA346766DD402A CRC64;
SQ
Query Match 43.9%; Score 134; DB 1; Length 192;
Best Local Similarity 46.0%; Pred. No. 2.1e-08;
Matches 29; Conservative 6; Mismatches 20; Indels 8; Gaps 1;
QY 1 YORPATAINNPVPTTYANPVPVPRHQAIPQRYLPNS-----HPPTVVRPNLH 52
DB 64 YQRPVALINNOFLPYPYAKPVAVRSPAQILQWVLPNTAPAKSCQDPTTMRHPHPH 123
QY 53 PSF 55
DB 124 LSF 126
RESULT 14
CASK_CAPCR STANDARD; PRT; 192 AA.
AC P42156;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein precursor.
GN CSN3 OR CSN10 OR CSNK.
OS Capricornis crispus (Japanese serow).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Capricornis.
OX NCBI_TaxID=9966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96139018; PubMed=8587130;
RA Chikuni K., Mori Y., Tabata T., Saito M., Monma M., Kosugiyama M.;
RT "Molecular phylogeny based on the kappa-casein and cytochrome b
RL J. Mol. Evol. 41:859-866(1995).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
CC casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14376; BAA03287.1;
DR EMBL; D14374; BAA03287.1; JOINED.
DR EMBL; D14375; BAA03287.1; JOINED.
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
DR Milk; Phosphorylation; Glycoprotein; Signal;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 192 KAPPA CASEIN.
FT SITE 126 127 CLEAVAGE (BY CHYMOSIN (-RENNIN))
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 152 152 O-LINKED (GALNAC... ) (BY SIMILARITY).
FT MOD_RES 172 172 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 189 189 PHOSPHORYLATION (BY SIMILARITY).
FT VARIANT 8 8 V->A.
FT SEQUENCE 192 AA; 21486 MW; 1261C8B4DBIAE9AC CRC64;
SQ
Query Match 42.3%; Score 129; DB 1; Length 192;
Best Local Similarity 44.4%; Pred. No. 7.8e-08;
Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;
QY 1 YORPATAINNPVPTTYANPVPVPRHQAIPQRYLPNS-----HPPTVVRPNLH 52
DB 64 YQRPVALINNOFLPYPYAKPVAVRSPAQILQWVLPNTAPAKSCQDPTTMRHPHPH 123
QY 53 PSF 55
DB 124 LSF 126
RESULT 15
CASK_CAPHI STANDARD; PRT; 192 AA.
AC P02670; Q8SPW9; Q8SPN0; Q8SPV1; Q8SPW6; Q8SPW7; Q8SPW8; Q8MWV5;
AC Q8WN74;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kappa casein precursor (Kappa-CN).
GN CSN3 OR CSN10 OR CSNK.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SSP. aegagrus; TISSUE=Mammary gland;
RX MEDLINE=94042560; PubMed=8226388;
RA Coll A., Folch J.M., Sanchez A.;
RT "Nucleotide sequence of the goat kappa-casein cDNA.";
RL J. Anim. Sci. 71:2833-2833(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96139018; PubMed=8587130;
RA Chikuni K., Mori Y., Tabata T., Saito M., Monma M., Kosugiyama M.;
RT "Molecular phylogeny based on the kappa-casein and cytochrome b
RL J. Mol. Evol. 41:859-866(1995).
RN [3]
RP SEQUENCE OF 22-192.
RA Mercier J.-C., Addeo F., Pellissier J.-P.;
RL Unpublished results, cited by:

```

RL Mercier J.-C., Chobert J.-M., Addeo F.;
RL FEBS Lett. 72:208-214(1976).
RN [4]
RP SEQUENCE OF 31-192 FROM N.A., AND VARIANTS ARG-65; ILE-86; ILE-140 AND
RP PRO-180.
RX MEDLINE-21422902; PubMed-11531704;
RA Caroli A., Jann O., Budelli E., Bolla P., Jaeger S., Erhardt G.;
RT "Genetic polymorphism of goat kappa-casein (CSN3) in different breeds
RT and characterization at DNA level";
RN Anim. Genet. 32:226-230(2001).
RN [5]
RP SEQUENCE OF 31-192 FROM N.A., AND VARIANTS ILE-86; ILE-140; VAL-177
RP AND PRO-180.
RX MEDLINE-21395138; PubMed-11504385;
RA Yahyaoui M.H., Coll A., Sanchez A., Folch J.M.;
RT "Genetic polymorphism of the caprine kappa casein gene";
RN J. Dairy Res. 68:209-216(2001).
RN [6]
RP SEQUENCE OF 31-192 FROM N.A., AND VARIANTS GLY-111 AND ILE-140.
RA Angiolillo A., Yahyaoui M.H., Sanchez A., Pilla F., Folch J.M.;
RT "Characterization of a new genetic variant in the caprine kappa-casein
RT gene";
RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 31-192 FROM N.A., AND VARIANTS ARG-65; ILE-86; ILE-140 AND
RP PRO-180.
RA Yahyaoui M.H., Sanchez A., Folch J.M.;
RT "Characterization of new genetic variants and genotyping of the
RT caprine kappa casein gene";
RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 31-189 FROM N.A., AND VARIANT ILE-140.
RA Fellugini M., Cubric-Curik V., Parma P., Curik I., Greppi G., Enne G.;
RT "Polymorphism of kappa-casein in Italian goat breeds: a new ACRS-PCR
RT designed DNA test for discrimination of A and B alleles";
RN Food Technol. Biotechnol. 40:293-298(2002).
RN [9]
RP SEQUENCE OF 127-192.
RX MEDLINE-77112689; PubMed-1016651;
RA Mercier J.-C., Addeo F., Pelissier J.-P.;
RT "Primary structure of the casein macropeptide of caprine kappa
RT casein";
RN Biochimie 58:1303-1310(1976).
CC -!- FUNCTION: Kappa casein stabilizes micelle formation, preventing
CC casein precipitation in milk.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: Belongs to the kappa-casein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14373; BAA03286.1;
DR EMBL; D14371; BAA03286.1; JOINED.
DR EMBL; D14372; BAA03286.1; JOINED.
DR EMBL; X60763; CA443174.1;
DR EMBL; AY027868; AAL17010.1;
DR EMBL; AF485339; AAL90871.1;
DR EMBL; AF485340; AAL90872.1;
DR EMBL; AF485341; AAL90873.1;
DR EMBL; AF485342; AAL90874.1;
DR EMBL; AF485343; AAL90875.1;
DR EMBL; AF485344; AAL90876.1;
DR EMBL; AF485345; AAL90877.1;
DR EMBL; AF485346; AAL90878.1;
DR EMBL; AF485347; AAL90879.1;
DR EMBL; AF485348; AAL90880.1;
DR EMBL; AF485349; AAL90881.1;
DR EMBL; AF485350; AAL90882.1;
DR EMBL; AF485351; AAL90883.1;
DR EMBL; AF485352; AAL90884.1;
DR EMBL; AF485353; AAL90885.1;
DR EMBL; AF485354; AAL90886.1;
DR EMBL; AF485355; AAL90887.1;
DR EMBL; AF485356; AAL90888.1;
DR EMBL; AF485357; AAL90889.1;
DR EMBL; AF485358; AAL90890.1;
DR EMBL; AF485359; AAL90891.1;
DR EMBL; AF485360; AAL90892.1;
DR EMBL; AF485361; AAL90893.1;
DR EMBL; AF485362; AAL90894.1;
DR EMBL; AF485363; AAL90895.1;
DR EMBL; AF485364; AAL90896.1;
DR EMBL; AF485365; AAL90897.1;
DR EMBL; AF485366; AAL90898.1;
DR EMBL; AF485367; AAL90899.1;
DR EMBL; AF485368; AAL90900.1;
DR EMBL; AF485369; AAL90901.1;
DR EMBL; AF485370; AAL90902.1;
DR EMBL; AF485371; AAL90903.1;
DR EMBL; AF485372; AAL90904.1;
DR EMBL; AF485373; AAL90905.1;
DR EMBL; AF485374; AAL90906.1;
DR EMBL; AF485375; AAL90907.1;
DR EMBL; AF485376; AAL90908.1;
DR EMBL; AF485377; AAL90909.1;
DR EMBL; AF485378; AAL90910.1;
DR EMBL; AF485379; AAL90911.1;
DR EMBL; AF485380; AAL90912.1;
DR EMBL; AF485381; AAL90913.1;
DR EMBL; AF485382; AAL90914.1;
DR EMBL; AF485383; AAL90915.1;
DR EMBL; AF485384; AAL90916.1;
DR EMBL; AF485385; AAL90917.1;
DR EMBL; AF485386; AAL90918.1;
DR EMBL; AF485387; AAL90919.1;
DR EMBL; AF485388; AAL90920.1;
DR EMBL; AF485389; AAL90921.1;
DR EMBL; AF485390; AAL90922.1;
DR EMBL; AF485391; AAL90923.1;
DR EMBL; AF485392; AAL90924.1;
DR EMBL; AF485393; AAL90925.1;
DR EMBL; AF485394; AAL90926.1;
DR EMBL; AF485395; AAL90927.1;
DR EMBL; AF485396; AAL90928.1;
DR EMBL; AF485397; AAL90929.1;
DR EMBL; AF485398; AAL90930.1;
DR EMBL; AF485399; AAL90931.1;
DR EMBL; AF485400; AAL90932.1;
DR EMBL; AF485401; AAL90933.1;
DR EMBL; AF485402; AAL90934.1;
DR EMBL; AF485403; AAL90935.1;
DR EMBL; AF485404; AAL90936.1;
DR EMBL; AF485405; AAL90937.1;
DR EMBL; AF485406; AAL90938.1;
DR EMBL; AF485407; AAL90939.1;
DR EMBL; AF485408; AAL90940.1;
DR EMBL; AF485409; AAL90941.1;
DR EMBL; AF485410; AAL90942.1;
DR EMBL; AF485411; AAL90943.1;
DR EMBL; AF485412; AAL90944.1;
DR EMBL; AF485413; AAL90945.1;
DR EMBL; AF485414; AAL90946.1;
DR EMBL; AF485415; AAL90947.1;
DR EMBL; AF485416; AAL90948.1;
DR EMBL; AF485417; AAL90949.1;
DR EMBL; AF485418; AAL90950.1;
DR EMBL; AF485419; AAL90951.1;
DR EMBL; AF485420; AAL90952.1;
DR EMBL; AF485421; AAL90953.1;
DR EMBL; AF485422; AAL90954.1;
DR EMBL; AF485423; AAL90955.1;
DR EMBL; AF485424; AAL90956.1;
DR EMBL; AF485425; AAL90957.1;
DR EMBL; AF485426; AAL90958.1;
DR EMBL; AF485427; AAL90959.1;
DR EMBL; AF485428; AAL90960.1;
DR EMBL; AF485429; AAL90961.1;
DR EMBL; AF485430; AAL90962.1;
DR EMBL; AF485431; AAL90963.1;
DR EMBL; AF485432; AAL90964.1;
DR EMBL; AF485433; AAL90965.1;
DR EMBL; AF485434; AAL90966.1;
DR EMBL; AF485435; AAL90967.1;
DR EMBL; AF485436; AAL90968.1;
DR EMBL; AF485437; AAL90969.1;
DR EMBL; AF485438; AAL90970.1;
DR EMBL; AF485439; AAL90971.1;
DR EMBL; AF485440; AAL90972.1;
DR EMBL; AF485441; AAL90973.1;
DR EMBL; AF485442; AAL90974.1;
DR EMBL; AF485443; AAL90975.1;
DR EMBL; AF485444; AAL90976.1;
DR EMBL; AF485445; AAL90977.1;
DR EMBL; AF485446; AAL90978.1;
DR EMBL; AF485447; AAL90979.1;
DR EMBL; AF485448; AAL90980.1;
DR EMBL; AF485449; AAL90981.1;
DR EMBL; AF485450; AAL90982.1;
DR EMBL; AF485451; AAL90983.1;
DR EMBL; AF485452; AAL90984.1;
DR EMBL; AF485453; AAL90985.1;
DR EMBL; AF485454; AAL90986.1;
DR EMBL; AF485455; AAL90987.1;
DR EMBL; AF485456; AAL90988.1;
DR EMBL; AF485457; AAL90989.1;
DR EMBL; AF485458; AAL90990.1;
DR EMBL; AF485459; AAL90991.1;
DR EMBL; AF485460; AAL90992.1;
DR EMBL; AF485461; AAL90993.1;
DR EMBL; AF485462; AAL90994.1;
DR EMBL; AF485463; AAL90995.1;
DR EMBL; AF485464; AAL90996.1;
DR EMBL; AF485465; AAL90997.1;
DR EMBL; AF485466; AAL90998.1;
DR EMBL; AF485467; AAL90999.1;
DR EMBL; AF485468; AAL91000.1;
DR EMBL; AF485469; AAL91001.1;
DR EMBL; AF485470; AAL91002.1;
DR EMBL; AF485471; AAL91003.1;
DR EMBL; AF485472; AAL91004.1;
DR EMBL; AF485473; AAL91005.1;
DR EMBL; AF485474; AAL91006.1;
DR EMBL; AF485475; AAL91007.1;
DR EMBL; AF485476; AAL91008.1;
DR EMBL; AF485477; AAL91009.1;
DR EMBL; AF485478; AAL91010.1;
DR EMBL; AF485479; AAL91011.1;
DR EMBL; AF485480; AAL91012.1;
DR EMBL; AF485481; AAL91013.1;
DR EMBL; AF485482; AAL91014.1;
DR EMBL; AF485483; AAL91015.1;
DR EMBL; AF485484; AAL91016.1;
DR EMBL; AF485485; AAL91017.1;
DR EMBL; AF485486; AAL91018.1;
DR EMBL; AF485487; AAL91019.1;
DR EMBL; AF485488; AAL91020.1;
DR EMBL; AF485489; AAL91021.1;
DR EMBL; AF485490; AAL91022.1;
DR EMBL; AF485491; AAL91023.1;
DR EMBL; AF485492; AAL91024.1;
DR EMBL; AF485493; AAL91025.1;
DR EMBL; AF485494; AAL91026.1;
DR EMBL; AF485495; AAL91027.1;
DR EMBL; AF485496; AAL91028.1;
DR EMBL; AF485497; AAL91029.1;
DR EMBL; AF485498; AAL91030.1;
DR EMBL; AF485499; AAL91031.1;
DR EMBL; AF485500; AAL91032.1;
DR EMBL; AF485501; AAL91033.1;
DR EMBL; AF485502; AAL91034.1;
DR EMBL; AF485503; AAL91035.1;
DR EMBL; AF485504; AAL91036.1;
DR EMBL; AF485505; AAL91037.1;
DR EMBL; AF485506; AAL91038.1;
DR EMBL; AF485507; AAL91039.1;
DR EMBL; AF485508; AAL91040.1;
DR EMBL; AF485509; AAL91041.1;
DR EMBL; AF485510; AAL91042.1;
DR EMBL; AF485511; AAL91043.1;
DR EMBL; AF485512; AAL91044.1;
DR EMBL; AF485513; AAL91045.1;
DR EMBL; AF485514; AAL91046.1;
DR EMBL; AF485515; AAL91047.1;
DR EMBL; AF485516; AAL91048.1;
DR EMBL; AF485517; AAL91049.1;
DR EMBL; AF485518; AAL91050.1;
DR EMBL; AF485519; AAL91051.1;
DR EMBL; AF485520; AAL91052.1;
DR EMBL; AF485521; AAL91053.1;
DR EMBL; AF485522; AAL91054.1;
DR EMBL; AF485523; AAL91055.1;
DR EMBL; AF485524; AAL91056.1;
DR EMBL; AF485525; AAL91057.1;
DR EMBL; AF485526; AAL91058.1;
DR EMBL; AF485527; AAL91059.1;
DR EMBL; AF485528; AAL91060.1;
DR EMBL; AF485529; AAL91061.1;
DR EMBL; AF485530; AAL91062.1;
DR EMBL; AF485531; AAL91063.1;
DR EMBL; AF485532; AAL91064.1;
DR EMBL; AF485533; AAL91065.1;
DR EMBL; AF485534; AAL91066.1;
DR EMBL; AF485535; AAL91067.1;
DR EMBL; AF485536; AAL91068.1;
DR EMBL; AF485537; AAL91069.1;
DR EMBL; AF485538; AAL91070.1;
DR EMBL; AF485539; AAL91071.1;
DR EMBL; AF485540; AAL91072.1;
DR EMBL; AF485541; AAL91073.1;
DR EMBL; AF485542; AAL91074.1;
DR EMBL; AF485543; AAL91075.1;
DR EMBL; AF485544; AAL91076.1;
DR EMBL; AF485545; AAL91077.1;
DR EMBL; AF485546; AAL91078.1;
DR EMBL; AF485547; AAL91079.1;
DR EMBL; AF485548; AAL91080.1;
DR EMBL; AF485549; AAL91081.1;
DR EMBL; AF485550; AAL91082.1;
DR EMBL; AF485551; AAL91083.1;
DR EMBL; AF485552; AAL91084.1;
DR EMBL; AF485553; AAL91085.1;
DR EMBL; AF485554; AAL91086.1;
DR EMBL; AF485555; AAL91087.1;
DR EMBL; AF485556; AAL91088.1;
DR EMBL; AF485557; AAL91089.1;
DR EMBL; AF485558; AAL91090.1;
DR EMBL; AF485559; AAL91091.1;
DR EMBL; AF485560; AAL91092.1;
DR EMBL; AF485561; AAL91093.1;
DR EMBL; AF485562; AAL91094.1;
DR EMBL; AF485563; AAL91095.1;
DR EMBL; AF485564; AAL91096.1;
DR EMBL; AF485565; AAL91097.1;
DR EMBL; AF485566; AAL91098.1;
DR EMBL; AF485567; AAL91099.1;
DR EMBL; AF485568; AAL91100.1;
DR EMBL; AF485569; AAL91101.1;
DR EMBL; AF485570; AAL91102.1;
DR EMBL; AF485571; AAL91103.1;
DR EMBL; AF485572; AAL91104.1;
DR EMBL; AF485573; AAL91105.1;
DR EMBL; AF485574; AAL91106.1;
DR EMBL; AF485575; AAL91107.1;
DR EMBL; AF485576; AAL91108.1;
DR EMBL; AF485577; AAL91109.1;
DR EMBL; AF485578; AAL91110.1;
DR EMBL; AF485579; AAL91111.1;
DR EMBL; AF485580; AAL91112.1;
DR EMBL; AF485581; AAL91113.1;
DR EMBL; AF485582; AAL91114.1;
DR EMBL; AF485583; AAL91115.1;
DR EMBL; AF485584; AAL91116.1;
DR EMBL; AF485585; AAL91117.1;
DR EMBL; AF485586; AAL91118.1;
DR EMBL; AF485587; AAL91119.1;
DR EMBL; AF485588; AAL91120.1;
DR EMBL; AF485589; AAL91121.1;
DR EMBL; AF485590; AAL91122.1;
DR EMBL; AF485591; AAL91123.1;
DR EMBL; AF485592; AAL91124.1;
DR EMBL; AF485593; AAL91125.1;
DR EMBL; AF485594; AAL91126.1;
DR EMBL; AF485595; AAL91127.1;
DR EMBL; AF485596; AAL91128.1;
DR EMBL; AF485597; AAL91129.1;
DR EMBL; AF485598; AAL91130.1;
DR EMBL; AF485599; AAL91131.1;
DR EMBL; AF485600; AAL91132.1;
DR EMBL; AF485601; AAL91133.1;
DR EMBL; AF485602; AAL91134.1;
DR EMBL; AF485603; AAL91135.1;
DR EMBL; AF485604; AAL91136.1;
DR EMBL; AF485605; AAL91137.1;
DR EMBL; AF485606; AAL91138.1;
DR EMBL; AF485607; AAL91139.1;
DR EMBL; AF485608; AAL91140.1;
DR EMBL; AF485609; AAL91141.1;
DR EMBL; AF485610; AAL91142.1;
DR EMBL; AF485611; AAL91143.1;
DR EMBL; AF485612; AAL91144.1;
DR EMBL; AF485613; AAL91145.1;
DR EMBL; AF485614; AAL91146.1;
DR EMBL; AF485615; AAL91147.1;
DR EMBL; AF485616; AAL91148.1;
DR EMBL; AF485617; AAL91149.1;
DR EMBL; AF485618; AAL91150.1;
DR EMBL; AF485619; AAL91151.1;
DR EMBL; AF485620; AAL91152.1;
DR EMBL; AF485621; AAL91153.1;
DR EMBL; AF485622; AAL91154.1;
DR EMBL; AF485623; AAL91155.1;
DR EMBL; AF485624; AAL91156.1;
DR EMBL; AF485625; AAL91157.1;
DR EMBL; AF485626; AAL91158.1;
DR EMBL; AF485627; AAL91159.1;
DR EMBL; AF485628; AAL91160.1;
DR EMBL; AF485629; AAL91161.1;
DR EMBL; AF485630; AAL91162.1;
DR EMBL; AF485631; AAL91163.1;
DR EMBL; AF485632; AAL91164.1;
DR EMBL; AF485633; AAL91165.1;
DR EMBL; AF485634; AAL91166.1;
DR EMBL; AF485635; AAL91167.1;
DR EMBL; AF485636; AAL91168.1;
DR EMBL; AF485637; AAL91169.1;
DR EMBL; AF485638; AAL91170.1;
DR EMBL; AF485639; AAL91171.1;
DR EMBL; AF485640; AAL91172.1;
DR EMBL; AF485641; AAL91173.1;
DR EMBL; AF485642; AAL91174.1;
DR EMBL; AF485643; AAL91175.1;
DR EMBL; AF485644; AAL91176.1;
DR EMBL; AF485645; AAL91177.1;
DR EMBL; AF485646; AAL91178.1;
DR EMBL; AF485647; AAL91179.1;
DR EMBL; AF485648; AAL91180.1;
DR EMBL; AF485649; AAL91181.1;
DR EMBL; AF485650; AAL91182.1;
DR EMBL; AF485651; AAL91183.1;
DR EMBL; AF485652; AAL91184.1;
DR EMBL; AF485653; AAL91185.1;
DR EMBL; AF485654; AAL91186.1;
DR EMBL; AF485655; AAL91187.1;
DR EMBL; AF485656; AAL91188.1;
DR EMBL; AF485657; AAL91189.1;
DR EMBL; AF485658; AAL91190.1;
DR EMBL; AF485659; AAL91191.1;
DR EMBL; AF485660; AAL91192.1;
DR EMBL; AF485661; AAL91193.1;
DR EMBL; AF485662; AAL91194.1;
DR EMBL; AF485663; AAL91195.1;
DR EMBL; AF485664; AAL91196.1;
DR EMBL; AF485665; AAL91197.1;
DR EMBL; AF485666; AAL91198.1;
DR EMBL; AF485667; AAL91199.1;
DR EMBL; AF485668; AAL91200.1;
DR EMBL; AF485669; AAL91201.1;
DR EMBL; AF485670; AAL91202.1;
DR EMBL; AF485671; AAL91203.1;
DR EMBL; AF485672; AAL91204.1;
DR EMBL; AF485673; AAL91205.1;
DR EMBL; AF485674; AAL91206.1;
DR EMBL; AF485675; AAL91207.1;
DR EMBL; AF485676; AAL91208.1;
DR EMBL; AF485677; AAL91209.1;
DR EMBL; AF485678; AAL91210.1;
DR EMBL; AF485679; AAL91211.1;
DR EMBL; AF485680; AAL91212.1;
DR EMBL; AF485681; AAL91213.1;
DR EMBL; AF485682; AAL91214.1;
DR EMBL; AF485683; AAL91215.1;
DR EMBL; AF485684; AAL91216.1;
DR EMBL; AF485685; AAL91217.1;
DR EMBL; AF485686; AAL91218.1;
DR EMBL; AF485687; AAL91219.1;
DR EMBL; AF485688; AAL91220.1;
DR EMBL; AF485689; AAL91221.1;
DR EMBL; AF485690; AAL91222.1;
DR EMBL; AF485691; AAL91223.1;
DR EMBL; AF485692; AAL91224.1;
DR EMBL; AF485693; AAL91225.1;
DR EMBL; AF485694; AAL91226.1;
DR EMBL; AF485695; AAL91227.1;
DR EMBL; AF485696; AAL91228.1;
DR EMBL; AF485697; AAL91229.1;
DR EMBL; AF485698; AAL91230.1;
DR EMBL; AF485699; AAL91231.1;
DR EMBL; AF485700; AAL91232.1;
DR EMBL; AF485701; AAL91233.1;
DR EMBL; AF485702; AAL91234.1;
DR EMBL; AF485703; AAL91235.1;
DR EMBL; AF485704; AAL91236.1;
DR EMBL; AF485705; AAL91237.1;
DR EMBL; AF485706; AAL91238.1;
DR EMBL; AF485707; AAL91239.1;
DR EMBL; AF485708; AAL91240.1;
DR EMBL; AF485709; AAL91241.1;
DR EMBL; AF485710; AAL91242.1;
DR EMBL; AF485711; AAL91243.1;
DR EMBL; AF485712; AAL91244.1;
DR EMBL; AF485713; AAL91245.1;
DR EMBL; AF485714; AAL91246.1;
DR EMBL; AF485715; AAL91247.1;
DR EMBL; AF485716; AAL91248.1;
DR EMBL; AF485717; AAL91249.1;
DR EMBL; AF485718; AAL91250.1;
DR EMBL; AF485719; AAL91251.1;
DR EMBL; AF485720; AAL91252.1;
DR EMBL; AF485721; AAL91253.1;
DR EMBL; AF485722; AAL91254.1;
DR EMBL; AF485723; AAL91255.1;
DR EMBL; AF485724; AAL91256.1;
DR EMBL; AF485725; AAL91257.1;
DR EMBL; AF485726; AAL91258.1;
DR EMBL; AF485727; AAL91259.1;
DR EMBL; AF485728; AAL91260.1;
DR EMBL; AF485729; AAL91261.1;
DR EMBL; AF485730; AAL91262.1;
DR EMBL; AF485731; AAL91263.1;
DR EMBL; AF485732; AAL91264.1;
DR EMBL; AF485733; AAL91265.1;
DR EMBL; AF485734; AAL91266.1;
DR EMBL; AF485735; AAL91267.1;
DR EMBL; AF485736; AAL91268.1;
DR EMBL; AF485737; AAL91269.1;
DR EMBL; AF485738; AAL91270.1;
DR EMBL; AF485739; AAL91271.1;
DR EMBL; AF485740; AAL91272.1;
DR EMBL; AF485741; AAL91273.1;
DR EMBL; AF485742; AAL91274.1;
DR EMBL; AF485743; AAL91275.1;
DR EMBL; AF485744; AAL91276.1;
DR EMBL; AF485745; AAL91277.1;
DR EMBL; AF485746; AAL91278.1;
DR EMBL; AF485747; AAL91279.1;
DR EMBL; AF485748; AAL91280.1;
DR EMBL; AF485749; AAL91281.1;
DR EMBL; AF485750; AAL91282.1;
DR EMBL; AF485751; AAL91283.1;
DR EMBL; AF485752; AAL91284.1;
DR EMBL; AF485753; AAL91285.1;
DR EMBL; AF485754; AAL91286.1;
DR EMBL; AF485755; AAL91287.1;
DR EMBL; AF485756; AAL91288.1;
DR EMBL; AF485757; AAL91289.1;
DR EMBL; AF485758; AAL91290.1;
DR EMBL; AF485759; AAL91291.1;
DR EMBL; AF485760; AAL91292.1;
DR EMBL; AF485761; AAL91293.1;
DR EMBL; AF485762; AAL91294.1;
DR EMBL; AF485763; AAL91295.1;
DR EMBL; AF485764; AAL91296.1;
DR EMBL; AF485765; AAL91297.1;
DR EMBL; AF485766; AAL91298.1;
DR EMBL; AF485767; AAL91299.1;
DR EMBL; AF485768; AAL91300.1;
DR EMBL; AF485769; AAL91301.1;
DR EMBL; AF485770; AAL91302.1;
DR EMBL; AF485771; AAL91303.1;
DR EMBL; AF485772; AAL91304.1;
DR EMBL; AF485773; AAL91305.1;
DR EMBL; AF485774; AAL91306.1;
DR EMBL; AF485775; AAL91307.1;
DR EMBL; AF485776; AAL91308.1;
DR EMBL; AF485777; AAL91309.1;
DR EMBL; AF485778; AAL91310.1;
DR EMBL; AF48


```

01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-DEC-2001 (TREMBlrel. 19, Last annotation update)
K-casein (Fragment).
Delphinidae gen. sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
unidentified Delphinidae.
NCBI_TaxID=49273;
[1]
SEQUENCE FROM N.A.
MEDLINE=96364219; PubMed=8752004;
Gatesy J., Hayashi C., Cronin M.A., Arcander P.:
"Evidence from milk casein genes that cetaceans are close relatives of
hippopotamid artiodactyls."
Mol Biol. Evol. 13:954-963(1996).
EMBL; U53687; AAB08407.1; -.
InterPro: IPR000117; Casein_kappa.
Pfam: PF00997; casein_kappa; 1.
ProDom: PD003689; Casein_kappa; 1.
NON_TER 1.
SEQUENCE 145 AA; 16483 MW; B02D33CA90EC38AD CRC64;

Query Match 63.0%; Score 192; DB 6; Length 145;
Best Local Similarity 61.8%; Pred. No. 5.5e-16;
Matches 34; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 YORPATAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSF 55
DQ 27 YQRPVALLNQMPYYPYAKVEVRPHQAIPQWQLPNIPTVAHHPHPRPSF 81

RESULT 3
P79230 PRELIMINARY; PRT; 145 AA.
AC P79230;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Kappa casein (Fragment).
OS Physeter catodon (Specim whale) (Physeter macrocephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
OC Physeteridae; Physeter.
OX NCBI_TaxID=9755;
[1]
SEQUENCE FROM N.A.
Gatesy J.;
"More DNA support for a Cetacea/Hippopotamidae clade: the blood
clotting protein gene g-fibrinogen.";
Mol. Biol. Evol. 0:0-0(1997).
EMBL; U86643; AAB47429.1; -.
InterPro: IPR000117; Casein_kappa.
Pfam: PF00997; casein_kappa; 1.
ProDom: PD003689; Casein_kappa; 1.
NON_TER 1.
SEQUENCE 145 AA; 16304 MW; EFF454A3A42BC3AB CRC64;

Query Match 61.3%; Score 187; DB 6; Length 145;
Best Local Similarity 60.0%; Pred. No. 2.3e-15;
Matches 33; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 YORPATAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSF 55
DQ 27 YQRPVALLNQMPYLYYAKPVVSPHAQIPQWQLPNIPTLAHHPHPRPSF 81

RESULT 4
Q8SPR0 PRELIMINARY; PRT; 185 AA.
AC Q8SPR0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

```

```

DE Kappa casein.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
[1]
SEQUENCE FROM N.A.
Lenasi T., Rogelj I., Dovc P.;
"Equus caballus kappa-casein (k-CN) mRNA."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY040863; AAK3669.1; -.
InterPro: IPR000117; Casein_kappa.
Pfam: PF00997; casein_kappa; 1.
ProDom: PD003689; Casein_kappa; 1.
NON_TER 1.
SEQUENCE 185 AA; 21021 MW; F7634F52CEE90606 CRC64;

Query Match 60.0%; Score 183; DB 6; Length 185;
Best Local Similarity 63.6%; Pred. No. 9.1e-15;
Matches 35; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

QY 1 YORPATAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSF 55
DQ 63 YQRLALLLNQHPYQYARPAVRPHQIPQWQLPNIPTVVRHPCPHPSF 117

RESULT 5
Q95MY5 PRELIMINARY; PRT; 90 AA.
AC Q95MY5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Kappa-casein (Fragment).
OS Connochaetes gnou (black wildebeest).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Alcelaphinae; Connochaetes.
OX NCBI_TaxID=59528;
[1]
SEQUENCE FROM N.A.
MEDLINE=21313697; PubMed=11420362;
Matthee C.A., Davis S.K.;
"Molecular Insights into the Evolution of the Family Bovidae: A
Nuclear DNA Perspective.";
RL Mol. Biol. Evol. 18:1220-1230(2001).
DR EMBL; AF210161; AAK67749.1; -.
InterPro: IPR000117; Casein_kappa.
Pfam: PF00997; casein_kappa; 1.
ProDom: PD003689; Casein_kappa; 1.
NON_TER 1.
NON_TER 90.
SEQUENCE 90 AA; 10592 MW; 4D164122E7C40DF1 CRC64;

Query Match 43.0%; Score 131; DB 6; Length 90;
Best Local Similarity 47.4%; Pred. No. 1e-08;
Matches 27; Conservative 5; Mismatches 17; Indels 8; Gaps 1;

QY 1 YORPATAINNPVPTTYANPAVVRPHQAIPQRYLPNSHPPTVVRPNLHPSF 49
DQ 34 YQRPVALLNQFLPYPIYAKPVAVRPPAQIQLQWQLPNTVPAKSCQDHPNTMARHP 90

RESULT 6
Q8WN75 PRELIMINARY; PRT; 159 AA.
AC Q8WN75;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Kappa casein (Fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

OC	Bovidae; Caprinae; Capra.
OX	NCBI_TaxID=9925;
[1]	
RN	SEQUENCE FROM N.A.
RA	Feligrini M., Cubric-Curik V., Curik I., Parma P., Greppi G., Enne G.;
RP	"Characterization of kappa casein alleles in the Italian goats by
RT	AIRS-PCR analysis."
RL	Food Technol. Biotechnol. 0:0-0(2002).
DR	EMBL; AF434987; AAL31534.1;
DR	InterPro; IPR000117; Casein_kappa.
DR	Pfam; PF00997; casein_kappa; 1.
DR	ProDom; PD003689; casein_kappa; 1.
FT	NON_TER 1
FT	NON_TER 159
SQ	SEQUENCE 159 AA; 17722 MW; 440E28AADF9FE8AB CRC64;
Query Match	
Best Local Similarity 42.3%; Score 129; DB 6; Length 159;	
Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;	
Qy	1 YQRPATAINNYPVPTYANPAVVRPHAQIPQOYLPS-----HPPTVVRRPNLH 52
Db	34 YQRPVALINNQLPYPYAKPVAVRSPAQTLQWQLPNTVPKSCDQPTTLARHPHPH 93
Qy	53 PSF 55
Db	94 LSF 96
RESULT 7	
Q8WN74	PRELIMINARY; PRT; 159 AA.
ID	Q8WN74
AC	Q8WN74; PRELIMINARY; PRT; 159 AA.
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Kappa casein (Fragment).
OS	Capra hircus (Goat).
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC	Rovidae; Caprinae; Capra.
OX	NCBI_TaxID=9925;
[1]	
RN	SEQUENCE FROM N.A.
RA	Feligrini M., Cubric-Curik V., Curik I., Parma P., Greppi G., Enne G.;
RP	"Characterization of kappa casein alleles in the Italian goats by
RT	AIRS-PCR analysis."
RL	Food Technol. Biotechnol. 0:0-0(2002).
DR	EMBL; AF434988; AAL31535.1;
DR	InterPro; IPR000117; Casein_kappa.
DR	Pfam; PF00997; casein_kappa; 1.
DR	ProDom; PD003689; Casein_kappa; 1.
FT	NON_TER 1
FT	NON_TER 159
SQ	SEQUENCE 159 AA; 17736 MW; 441923C77E4FE8AB CRC64;
Query Match	
Best Local Similarity 42.3%; Score 129; DB 6; Length 159;	
Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;	
Qy	1 YQRPATAINNYPVPTYANPAVVRPHAQIPQOYLPS-----HPPTVVRRPNLH 52
Db	34 YQRPVALINNQLPYPYAKPVAVRSPAQTLQWQLPNTVPKSCDQPTTLARHPHPH 93
Qy	53 PSF 55
Db	94 LSF 96
RESULT 8	
Q8SPNO	PRELIMINARY; PRT; 162 AA.
ID	Q8SPNO
AC	Q8SPNO;

```
QY 53 PSF 55
Db 94 LSF 96

RESULT 10
Q8SPW8 PRELIMINARY; PRT; 162 AA.
ID Q8SPW8
AC Q8SPW8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Kappa casein (Fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21395138; PubMed=11504385;
RA Yahyaoui M.H., Coll A., Sanchez A., Folch J.M.;
RT "Genetic polymorphism of the caprine kappa casein gene.";
RL J. Dairy Res. 68:209-216(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Yahyaoui M.H., Coll A., Folch J.M., Sanchez A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF485339; AAL90871.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
FT NON_TER 1
SQ SEQUENCE 162 AA; 18051 MW; B8D460D40E28AADF CRC64;

Query Match 42.3%; Score 129; DB 6; Length 162;
Best Local Similarity 44.4%; Pred. No. 3.5e-08;
Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;

QY 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQOYLPNS-----HPPTVVRPNLH 52
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34 YQORPVALINNOFLPYPIYAKPVAVRSQAQTLQWQLVNTVPKSCQDQPTTLARHPH 93

QY 53 PSF 55
Db 94 LSF 96

RESULT 11
Q8SPV1 PRELIMINARY; PRT; 162 AA.
ID Q8SPV1
AC Q8SPV1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Kappa casein (Fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Angiolillo A., Yahyaoui M.H., Sanchez A., Pilla F., Folch J.M.;
RT "Characterization of a new genetic variant in the caprine k-casein gene.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486523; AAL93193.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.

QY 53 PSF 55
Db 94 LSF 96

RESULT 12
Q8SPW9 PRELIMINARY; PRT; 162 AA.
ID Q8SPW9
AC Q8SPW9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Kappa casein (Fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Yahyaoui M.H., Sanchez A., Folch J.M.;
RT "Characterization of new genetic variants and genotyping of the caprine kappa casein gene.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY090467; RAM12027.1;
DR InterPro; IPR000117; Casein_kappa.
DR Pfam; PF00997; casein_kappa; 1.
DR ProDom; PD003689; Casein_kappa; 1.
FT NON_TER 1
SQ SEQUENCE 162 AA; 18089 MW; 3BA5A9E590C465F4 CRC64;

Query Match 42.0%; Score 128; DB 6; Length 162;
Best Local Similarity 44.4%; Pred. No. 4.6e-08;
Matches 28; Conservative 6; Mismatches 21; Indels 8; Gaps 1;

QY 1 YORRPAIAINNPVPTTYANPAVVRPHQAIPQOYLPNS-----HPPTVVRPNLH 52
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34 YQORPVALINNOFLPYPIYAKPIAVRSQAQTLQWQLVNTVPKSCQDQPTTLARHPH 93

QY 53 PSF 55
Db 94 LSF 96

RESULT 13
Q8SPW6 PRELIMINARY; PRT; 162 AA.
ID Q8SPW6
AC Q8SPW6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Kappa casein (Fragment).
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21395138; PubMed=11504385;
RA Yahyaoui M.H., Coll A., Sanchez A., Folch J.M.;
```

```

RESULT 15
Q9N273
ID Q9N273 PRELIMINARY; PRT; 160 AA.
AC Q9N273;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Kappa-casein (Fragment).
GN CSN3.

```

Search completed: August 11, 2003, 08:20:51
Job time : 34.0388 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein -- protein search, using sw model

Run on: August 11, 2003; 08:16:37 ; Search time 34.9515 Seconds
(without alignments)
217.984 Million cell updates/sec

Title: US-09-508-095-19

Perfect score: 267

Sequence: 1 GRRSVQWCAVSQPEATKCF.....VRGPPVSCIKRDSPIQCIQA 48

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	100.0	48	AAW93883	Bifidobacterium bi
2	261	97.8	49	AAW93884	Bifidobacterium bi
3	261	97.8	51	AAW93885	Human lactoferrin
4	261	97.8	52	AAW93886	Human lactoferrin
5	261	97.8	52	AAW93887	Lactoferrin decomp
6	261	97.8	54	AAW93888	Amino acid sequenc
7	261	97.8	690	AAW93889	Human lactoferrin
8	261	97.8	690	AAW93890	Human codon optimi
9	261	97.8	690	AAW93891	Codon optimised la

10	261	97.8	692	21	AAW58733	Human lactoferrin.
11	261	97.8	692	22	AAW77906	Human lactoferrin.
12	261	97.8	692	23	AAW97382	Human lactoferrin
13	261	97.8	692	23	AAW97382	Human codon optimi
14	261	97.8	693	16	AAW85146	Lactoferrin protei
15	261	97.8	705	13	AAW85146	Human lactoferrin
16	261	97.8	705	20	AAW31152	Human lactoferrin
17	261	97.8	708	13	AAW22424	Human lactoferrin
18	261	97.8	708	20	AAW31153	Human lactoferrin
19	261	97.8	709	15	AAW45199	Human lactoferrin.
20	261	97.8	709	15	AAW45199	Human lactoferrin.
21	261	97.8	711	19	AAW53880	Bovine alphaSI-cas
22	261	97.8	711	21	AAW53881	Human lactoferrin
23	261	97.8	711	21	AAW77578	Chronic hepatitis
24	261	97.8	711	22	AAW64828	Human lactoferrin
25	261	97.8	711	22	AAW64828	Human lactoferrin
26	256	95.9	689	22	AAW77909	Lactoferrin 6 kD v
27	255	95.5	51	21	AAW22824	Lactoferrin 6 kD v
28	254	95.1	49	20	AAW93882	Bifidobacterium bi
29	254	95.1	51	21	AAW22822	Lactoferrin 6 kD v
30	254	95.1	51	21	AAW22823	Lactoferrin 6 kD v
31	254	95.1	690	19	AAW71180	Mutant human lacto
32	254	95.1	709	18	AAW21695	Human lactoferrin
33	254	95.1	709	19	AAW53879	Human lactoferrin.
34	254	95.1	709	21	AAW38330	Human lactoferrin.
35	254	95.1	709	21	AAW77577	Human lactoferrin.
36	254	95.1	709	22	AAW6842	Protein encoded by
37	254	95.1	711	11	AAW80833	Human lactoferrin.
38	254	95.1	711	14	AAW43653	Lactoferrin. Homo
39	254	95.1	711	17	AAW09342	Human lactoferrin.
40	254	95.1	711	19	AAW57317	Human lactoferrin.
41	254	95.1	711	20	AAW86021	Amino acid sequenc
42	254	95.1	711	21	AAW808182	Human lactoferrin
43	254	95.1	711	22	AAW02341	Lactoferrin derive
44	252	94.4	47	17	AAW10519	Antifungal peptide
45	252	94.4	47	18	AAW24270	

ALIGNMENTS

RESULT 1
AAW93883
ID AAW93883 standard; peptide; 48 AA.
XX
XX AAW93883;
XX
XX
DT 25-JUN-1999 (first entry)
XX
DE Bifidobacterium bifidus stimulating peptide 19.
XX

XX Bifidogenic peptide; protease; treatment; microbe-related disease;
KW bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;
KW infection; inflammation; microbial induced tumour; degenerative disorder;
KW diarrhoea; colic; oral microflora; intestinal microflora; caries;
KW vaginal microflora.
XX
XX Bifidobacterium bifidus.
OS

Key Location/Qualifiers
FH Disulfide-bond 9.45
FT Disulfide-bond 19..36
XX
XX W09914231-A2.
XX
XX 25-MAR-1999.
XX
XX 16-SEP-1998; 98WO-EP05899.
XX
XX 11-FEB-1998; 98DE-1005385.
XX 16-SEP-1997; 97DE-1040604.
XX
XX (FORS/) FORSMANN W.

XX PI Forssmann W, Liepke C, Zucht H;
 XX DR WPI; 1999-244022/20.
 XX PT Milk-derived peptides that stimulate Bifidobacterium bifidus
 XX PS Claim 2; Page 3; 25pp; German.
 XX CC This invention describes milk-derived bifidogenic peptides and their
 CC active derivatives or fragments, and combinations of them produced by
 CC chemical coupling. Such are produced from bovine or human milk by
 CC treatment for 2 hr with proteases, then centrifuging to remove fat and
 CC acidifying to pH 2 to precipitate proteins. The solution phase is then
 CC subjected to reverse-phase high-performance liquid chromatography (HPLC)
 CC and cation-exchange HPLC, the fractions adjusted to salt content below
 CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by
 CC growing Bifidobacterium bifidus and Escherichia coli in presence of the
 CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15
 CC are selected where Bw = germ count after 16 hr culture of B. bifidus in
 CC similar conditions in a peptide-free control, Ew = germ count under
 CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at
 CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free
 CC control. The peptides AAW93865-W93888 are used to treat microbe-related
 CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,
 CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,
 CC microbially induced tumours or degenerative disorders, diarrhoea, colic,
 CC abnormalities in oral, intestinal or vaginal microflora, or caries.
 XX SQ Sequence 48 AA;

Query Match 100.0%; Score 267; DB 20; Length 48;
 Best Local Similarity 100.0%; Pred. No. 2.8e-29;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRRSVQWCAVSQPEATKCFQWQNNMKYRGPPVSCIKRDSPIQCIQA 48
 DB 1 GRRSVQWCAVSQPEATKCFQWQNNMKYRGPPVSCIKRDSPIQCIQA 48

RESULT 2
 AAW93884
 ID AAW93884 standard; peptide; 49 AA.
 AC AAW93884;

XX DT 25-JUN-1999 (first entry)
 XX DE Bifidobacterium bifidus stimulating peptide 20.

XX KW Bifidogenic peptide; protease; treatment; microbe-related disease;
 KW bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;
 KW infection; inflammation; microbial induced tumour; degenerative disorder;
 KW diarrhoea; colic; oral microflora; intestinal microflora; caries;
 XX KW vaginal microflora.

XX OS Bifidobacterium bifidus.

XX FH Key Location/Qualifiers
 FT Disulfide-bond 10..46
 FT Disulfide-bond 20..37

XX PN W09914231-A2.

XX PD 25-MAR-1999.

XX PF 16-SEP-1998; 98WO-EP05899.

XX PR 11-FEB-1998; 98DE-1005385.

XX PR 16-SEP-1997; 97DE-1040604.

XX PA (FORS/) FORSSMANN W.

XX PI Forssmann W, Liepke C, Zucht H;
 XX DR WPI; 1999-244022/20.
 XX PT Milk-derived peptides that stimulate Bifidobacterium bifidus
 XX PS Claim 2; Page 3; 25pp; German.

XX CC This invention describes milk-derived bifidogenic peptides and their
 CC active derivatives or fragments, and combinations of them produced by
 CC chemical coupling. Such are produced from bovine or human milk by
 CC treatment for 2 hr with proteases, then centrifuging to remove fat and
 CC acidifying to pH 2 to precipitate proteins. The solution phase is then
 CC subjected to reverse-phase high-performance liquid chromatography (HPLC)
 CC and cation-exchange HPLC, the fractions adjusted to salt content below
 CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by
 CC growing Bifidobacterium bifidus and Escherichia coli in presence of the
 CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15
 CC are selected where Bw = germ count after 16 hr culture of B. bifidus in
 CC similar conditions in a peptide-free control, Ew = germ count under
 CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at
 CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free
 CC control. The peptides AAW93865-W93888 are used to treat microbe-related
 CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,
 CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,
 CC microbially induced tumours or degenerative disorders, diarrhoea, colic,
 CC abnormalities in oral, intestinal or vaginal microflora, or caries.

XX SQ Sequence 49 AA;

Query Match 97.8%; Score 261; DB 20; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.9e-28;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRSVQWCAVSQPEATKCFQWQNNMKYRGPPVSCIKRDSPIQCIQA 48
 DB 3 RRRSVQWCAVSQPEATKCFQWQNNMKYRGPPVSCIKRDSPIQCIQA 49

RESULT 3
 AAB22817
 ID AAB22817 standard; protein; 51 AA.
 AC AAB22817;

XX DT 03-JAN-2001 (first entry)
 XX DE Human lactoferrin fragment, SEQ ID NO:2.

XX KW Human lactoferrin fragment; host-defence peptide; antimicrobial;
 KW endotoxin neutralisation; Gram negative bacterium; lipopolysaccharide;
 KW LPS; antibacterial; antifungal; sepsis; septic shock; tuberculosis;
 KW leprosy.

XX OS Homo sapiens.

XX PN W0200049040-A2.

XX PD 24-AUG-2000.

XX PF 27-JAN-2000; 2000WO-IB00271.

XX PR 05-FEB-1999; 99US-0245527.

XX PA (ENDO-) ENDOGEN RES PH AB.

XX PI Mann DM;

XX DR WPI; 2000-558291/51.

XX PT Antimicrobial/endotoxin neutralizing polypeptide generated by

PT proteolytic digestion of lactoferrin, useful for treatment of bacterial
 PT and fungal infections, and for preventing contamination of e.g. food
 PT products, living cells, and blood

XX Claim 5; Page 8; 90pp; English.

XX The invention relates to proteolytic fragments of lactoferrin (LF)
 CC and variants thereof which act as host-defence peptides, having
 CC antimicrobial and endotoxin-neutralising activity. The peptides are
 CC useful for treatment of bacterial and fungal infections, and may be
 CC particularly used for treating sepsis, tuberculosis and leprosy. The
 CC peptides are also useful for treating a variety of products at risk of
 CC contamination with microorganisms and endotoxin (lipopolysaccharide,
 CC LPS), such as food products, tissue, living cells, blood, drugs,
 CC glassware and surgical equipment. Sequences AAR22817 and AAR22822-B22827
 CC represent specifically claimed human lactoferrin-derived polypeptides of
 CC the invention. The present sequence is a 6 kD fragment of human
 CC lactoferrin comprising the N-terminal 51 residues produced via cathepsin
 CC D digestion.

XX Sequence 51 AA;

Query Match 97.8%; Score 261; DB 21; Length 51;
 Best Local Similarity 100.0%; Pred. No. 2e-28;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 48
 |||||
 DB 3 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 49

RESULT 4

AAR87906
 ID AAR87906 standard; peptide; 52 AA.

XX AAR87906;

XX 01-MAR-1996 (first entry)

XX Human lactoferrin (1-52).

XX antiviral; lactoferrin;

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 10..46

FT /note= "each Cys residue may be in reduced form or may
 form a disulphide bond with another Cys residue.
 A 10-46 disulphide bond is exemplified"

FT Disulfide-bond 20..37

FT /note= "each Cys residue may be in reduced form or may
 form a disulphide bond with another Cys residue.
 A 20-37 disulphide bond is exemplified"

FT JP07069915-A.

XX 14-MAR-1995.

XX 02-SEP-1993; 93JP-0240284.

XX 02-SEP-1993; 93JP-0240284.

XX (SNOW) SNOW BRAND MILK PROD CO LTD.

XX WPI; 1995-144726/19.

XX Inhibitor against viral infection and proliferation - contains
 PT peptide having sequence from lactoferrin

XX Claim 2; Page 2; 10pp; Japanese.

XX The sequence is one of six peptides disclosed as having inhibitory

CC effect against viral infection. The peptides are derived from
 CC lactoferrin. Their activity is demonstrated against cytomegalovirus.
 XX
 SQ Sequence 52 AA;

Query Match 97.8%; Score 261; DB 16; Length 52;
 Best Local Similarity 100.0%; Pred. No. 2e-28;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 48
 |||||

DB 3 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 49

RESULT 5

AAR91192

ID AAR91192 standard; peptide; 52 AA.

XX AAR91192;

XX 06-SEP-1996 (first entry)

XX Lactoferrin decomposition peptide.

XX immunoactivator; antiviral; cytomegalovirus; cosmetic; food; feed;

XX lymphocyte blastogenesis.

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 10..46

FT /note= "optionally this bond may be reduced"

FT Disulfide-bond 20..37

FT /note= "optionally this bond may be reduced"

XX JP08073499-A.

XX 19-MAR-1996.

XX 01-SEP-1994; 94JP-0232026.

XX 01-SEP-1994; 94JP-0232026.

XX (SNOW) SNOW BRAND MILK PROD CO LTD.

XX WPI; 1996-205535/21.

XX New peptide(s) derived from human lactoferrin - are useful as
 PT immuno-activators, esp. for preventing infection by cytomegalovirus

XX Claim 3; Page 2; lipp; Japanese.

XX The present sequence is one of four new peptides (see AAR91191 -
 CC AAR91193) obtained by enzymatically decomposing lactoferrin using
 CC protease. The peptides are immunoactivators which induce
 CC blastogenesis of lymphocytes and are particularly useful for
 CC treating cytomegalovirus infection. They can be used in drugs and
 CC cosmetics and can be added to foods and feeds.

XX Sequence 52 AA;

Query Match 97.8%; Score 261; DB 17; Length 52;
 Best Local Similarity 100.0%; Pred. No. 2e-28;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 48
 |||||

DB 3 RRRSVQWCAVSQPEATKCFQWRNMKVRGPPVSCIKRDSPIQICQA 49

RESULT 6

AAR88218

ID AAR88218 standard; peptide; 54 AA.

```

XX AAR88218;
XX
XX 12-JUN-1996 (first entry)
XX
XX Amino acid sequence contg. lactoferrisin antibacterial peptide.
XX
XX Recombinant vector; human; lactoferrisin; antibacterial peptide;
XX regulatory sequence; tac; promoter; shuttle; vector; pGEX2; GAL1;
XX pK0M2; Rous Sarcoma Virus; long terminal repeat; pRSVNot;
XX lactoferricin.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 19..29
XX /note= "lactoferrisin antibacterial peptide"
XX
XX JP07274970-A.
XX
XX 24-OCT-1995.
XX
XX 01-APR-1994; 94JP-0085244.
XX
XX 01-APR-1994; 94JP-0085244.
XX
XX (MORG ) MORINAGA MILK IND CO LTD.
XX
XX WPI; 1995-399338/51.
XX
XX Recombinant vector contg. lactoferrisin gene - used to prepare an
XX antibacterial peptide
XX
XX Claim 7; Page 11; 18pp; Japanese.
XX
XX A recombinant vector in which a DNA sequence encoding at least the
XX generic lactoferrisin antibacterial peptide AAR88216 (specific
XX examples of which are given in AAR84083-85) is inserted, pref.
XX downstream of the vector's regulatory sequence, is claimed. Pref.
XX examples of amino acid sequences contg. the above peptide are
XX given in AAR88217/18. The regulatory sequence is the tac promoter
XX from shuttle vector pGEX2, the GAL1 promoter from vector pK0M2 or
XX Rous Sarcoma Virus long terminal repeat from vector pRSVNot.
XX
XX Sequence 54 AA;
XX
XX Query Match 97.8%; Score 261; DB 16; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-28;
XX Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 RRRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQICQA 48
XX
XX 3 RRRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQICQA 49
XX
XX
XX RESULT 7
XX AAG77908
XX ID AAG77908 standard; protein; 690 AA.
XX
XX AAG77908;
XX
XX 22-JAN-2002 (first entry)
XX
XX Human lactoferrin variant hLF-2N.
XX
XX Human; lactoferrin; hLF; infectious disease; inflammatory disease;
XX excess of heparin; gastroenteritis; inflammatory bowel disease;
XX sepsis; anaemia; myelopoiesis; reperfusion injury; cytokine release;
XX proteoglycan; hLF-2N; mutant; mutein.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers

```

```

FT Domain 26..29
XX /label= Cationic_domain
XX
XX WO200172322-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-NL00253.
XX
XX 27-MAR-2000; 2000EP-0201110.
XX
XX 27-MAR-2000; 2000US-193352P.
XX
XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.
XX
XX Van Bree JBMM, Nuijens JH;
XX
XX WPI; 2001-648424/74.
XX
XX Use of lactoferrin for treatment of infectious diseases, inflammatory
XX diseases and excess of heparin
XX
XX Claim 10; Page -; 49pp; English.
XX
XX The sequence represents the human lactoferrin (hLF) protein variant
XX hLF-2N. The invention relates to novel methods of treatment using high
XX doses of lactoferrin. The methods of the invention are useful for the
XX treatment or prophylaxis of infectious diseases, inflammatory diseases
XX and excess of heparin e.g. gastroenteritis, inflammatory bowel disease,
XX sepsis, anaemia, myelopoiesis, reducing reperfusion injury, cytokine
XX release and proteoglycan-mediated entry of virus into cells. The
XX advantage of the method is that the patient is substantially free of side
XX effect responses to administration of lactoferrin. Therefore large doses
XX of lactoferrin can be administered.
XX
XX Note: The present sequence is not shown in the specification but is
XX derived from human wild-type lactoferrin sequence given in the sequence
XX listing of the specification.
XX
XX Sequence 690 AA;
XX
XX Query Match 97.8%; Score 261; DB 22; Length 690;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-27;
XX Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 RRRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQICQA 48
XX
XX 1 RRRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQICQA 47
XX
XX
XX RESULT 8
XX AAE27884
XX ID AAE27884 standard; Protein; 690 AA.
XX
XX AAE27884;
XX
XX 27-DEC-2002 (first entry)
XX
XX Human codon optimised lactoferrin.
XX
XX Human; feed improvement; plant-derived feed; antibiotic; additive;
XX anti-microbial; poultry; lactoferrin; flour; malt.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 319..320
XX /note= "Encoded by CTG TAC CTC"
XX
XX WO200263975-A2.
XX
XX 22-AUG-2002.
XX
XX 14-FEB-2002; 2002WO-US04919.
XX
XX

```



```
XX 14-FEB-2001; 2001US-269188P.
PR 02-MAY-2001; 2001US-0847232.
XX (VENT-) VENTRIA BIOSCIENCE.
XX Huang N, Rodriguez RL, Hagie FE;
XX WPI; 2002-582708/73.
DR N-PSDB; AAB45297.
XX Improved feed for production animals, comprising plant-derived feed
PT ingredients, and seed composition containing flour, extract, or malt
PT from mature monocot seeds and heterologous seed-produced anti-microbial
PT proteins
XX Example 7; Page 148-150; 175pp; English.
XX The invention relates to improved feed for production animals,
CC comprising one or more plant-derived feed ingredients, substantially
CC unsupplemented with small-molecule antibiotics and as an additive a
CC seed composition containing a flour, extract or malt obtained from
CC mature monocot seeds and one or more heterologous seed-produced anti-
CC microbial proteins in substantially unpurified form. The invention
CC is useful as a feed for production animals such as poultry and hoofed
CC farm animals. The present sequence is human codon optimised lactoferrin.
CC This sequence is used in the invention.
XX Sequence 690 AA;
SQ Query Match 97.8%; Score 261; DB 23; Length 690;
Best Local Similarity 100.0%; Pred. No. 3.7e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQCIQA 48
Db 3 RRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQCIQA 49
RESULT 9
ABG80724
ID ABG80724 standard; Protein; 690 AA.
XX AC ABG80724;
XX 29-NOV-2002 (first entry)
XX Codon optimised lactoferrin protein.
XX Transformed plant; heterologous transcription factor; transgenic plant;
XX seed protein; protein expression.
XX Homo sapiens.
XX WO200264750-A2.
XX 22-AUG-2002.
XX 14-FEB-2002; 2002WO-US04909.
XX 14-FEB-2001; 2001US-269188P.
XX 14-FEB-2001; 2001US-269199P.
XX 02-MAY-2001; 2001US-0847232.
XX (VENT-) VENTRIA BIOSCIENCE.
XX Huang N, Yang D;
XX WPI; 2002-657592/70.
DR N-PSDB; ABS66515.
XX Producing heterologous polypeptide in plant grain, by culturing
PT transformed plant to form a grain-producing transforming plant, and
```

```
PT recovering transgenic grains containing polypeptide from transformed
XX plant
PS Example 15; Page 117; 230pp; English.
XX The invention describes a method of producing a heterologous polypeptide
CC (I) in a grain of a plant, comprising culturing a transformed plant (P1)
CC comprising a first chimeric gene, and optionally, at least one
CC heterologous transcription factor that is capable of enhancing the
CC expression of the chimeric gene, to form a grain producing transforming
CC plant (P2), and recovering transgenic grains containing (I) from P2.
CC The method is useful for producing heterologous polypeptide in a grain
CC of a plant. (I) is a non-plant storage, human or non-human animal, milk
CC or other than a milk polypeptide, antibodies, cytokines, lymphokines,
CC chemokines, hormones, growth factors, coagulation factors,
CC anti-infectives, or cytotoxins, or anti-inflammatory molecule or
CC intestinal trefoil factor (ITF) or its active fragment. Preferably, (I)
CC is lactoferrin, lysozyme, lactoferricin, ITF, epidermal growth factor
CC (EGF), keratinocyte growth factor (KGF), insulin-like growth factor I
CC (IGF-I), lactoferrin, kappa-casein, haptocorrin, lactoperoxidase,
CC alpha-1-antitrypsin, immunoglobulins, alpha-lactalbumin,
CC beta-lactoglobulin, alpha-casein, beta-casein, albumin, fibrinogen or
CC protease inhibitor. This is the amino acid sequence of a protein
CC associated with method of producing a transgenic plant.
XX Sequence 690 AA;
SQ Query Match 97.8%; Score 261; DB 23; Length 690;
Best Local Similarity 100.0%; Pred. No. 3.7e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQCIQA 48
Db 3 RRSVQWCAVSQPEATKCFQWQRMNMRKVRGPPVSCIKRDSPIQCIQA 49
RESULT 10
AAV58733
ID AAV58733 standard; Protein; 692 AA.
XX AC AAV58733;
XX 25-APR-2000 (first entry)
XX Human lactoferrin.
XX Lactoferrin; human; transgenic plant.
XX Homo sapiens.
XX WO200004146-A1.
XX 27-JAN-2000.
XX 19-JUL-1999; 99WO-IT00226.
XX 17-JUL-1998; 98IT-RM00478.
XX (PLAN-) PLANTECHNO SRL.
XX Fogher C;
XX WPI; 2000-161129/14.
DR N-PSDB; AAZ58122.
XX Synthetic polynucleotide encoding human lactoferrin, used for
PT production of functional foods, vegetal milks and human lactoferrin
XX Disclosure; Page 73-77; 93pp; English.
XX The present sequence is that of human lactoferrin. The invention
CC relates to a synthetic gene (see AAV58122) that encodes human
CC lactoferrin but which has codon usage designed to maximise
```

Query Match	97.8%	Score 261;	DB 22;	Length 692;
Best Local Similarity	100.0%;	Pred. No. 3.7e-27;		
Matches 47;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2	RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQICQA	48	
Db	3	RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQICQA	49	
RESULT 12				
AAB97382				
ID	AAB97382 standard; protein: 692 AA.			
XX	AC	AAB97382;		
XX	XX	17-AUG-2001 (first entry)		
DT	XX	Human lactoferrin (hLF).		
DE	XX	Human; lactoferrin; hLF; N-terminal; antimicrobial; heparin;		
XX	XX	inflammatory response; cytokine production reduction;		
KW	KW	neutrophil degranulation; myelopoiesis inhibition.		
KW	KW	Homo sapiens.		
OS	XX			
XX	XX	Key		
FF	FF	Location/Qualifiers		
FT	FT	1..27		
FT	FT	/note= "Fragments of the N-terminal are specifically referred to in the claims"		
FT	FT	2..5		
FT	FT	/label= Cationic_domain		
FT	FT	28..31		
FT	FT	/label= Cationic_domain		
FT	FT	WO200134641-A2.		
XX	XX	17-MAY-2001.		
XX	XX	10-NOV-2000; 2000WO-NL00821.		
XX	XX	11-NOV-1999; 99EP-0203775.		
PR	PR	11-NOV-1999; 99US-0164975.		
XX	XX	(PHAR-) PHARMING INTELLECTUAL PROPERTY BV.		
PA	XX	Van Berkel PHC, Nibbering PH, Nuijens JH;		
XX	XX	WPI; 2001-335909/35.		
DR	XX	New polypeptides comprising the N-terminal region of human lactoferrin		
PT	PT	protein exhibit higher antimicrobial activity than the full length		
PT	PT	protein and are useful to treat bacterial infections -		
XX	XX	Claim 1; Page 55-57; 59pp; English.		
XX	XX	This invention relates to fragments (between 6 and 26 amino acids) of the		
XX	XX	human lactoferrin hLF protein (represented by the present sequence)		
CC	CC	N-terminal hLF peptides have antimicrobial activity. The peptides of the		
CC	CC	invention are used to treat microbial infections, especially infections		
CC	CC	by gram positive or negative bacteria, particularly Listeria,		
CC	CC	Staphylococcus, Klebsiella or Escherichia species, especially		
CC	CC	L. monocytogenes, S. aureus and E. coli. Other uses include reducing		
CC	CC	inflammatory response by neutralising heparin or lipopolysaccharide or by		
CC	CC	reducing cytokine production and neutrophil degranulation, inhibiting		
CC	CC	entry of viruses such as cytomegalovirus, herpes simplex virus 1 or HIV		
CC	CC	into cells, inhibiting myelopoiesis and reducing production of GM-CSF		
XX	XX	(granulocyte/macrophage colony stimulating factor).		
XX	XX	Sequence 692 AA;		
SQ	SQ	Query Match 97.8%; Score 261; DB 22; Length 692;		

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 48
 |||||
 Db 3 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 49

RESULT 13

AAE28002
 ID AAE28002 standard; Protein; 692 AA.

XX AAE28002;

XX 13-DEC-2002 (first entry)

XX Human codon optimised lactoferrin.

XX Human; plant-derived food; flour; malt; monocot seed; milk protein;
 KW infant formula; lactoferrin.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 320..321

XX /note= "These amino acid residues are absent in the
 sequence shown as SEQ ID NO:2 in sequence listing of
 the specification"

XX WO200264814-A2.

XX 22-AUG-2002.

XX 14-FEB-2002; 2002WO-US04921.

XX 14-FEB-2001; 2001US-269199P.

XX 02-MAY-2001; 2001US-0847232.

XX (VENT-) VENTRIA BIOSCIENCE.

XX Huang N, Rodriguez RL, Hagie FE;

XX WPI; 2002-667011/71.

XX N-PSDB; AAD44922.

XX New nutritionally enhanced food compositions, useful for improving
 infant formula, or supplementing or enhancing the diet of infants,
 particularly very-low birth weight infants

XX Example 9; Page 104; 179pp; English.

XX The invention relates to a nutritionally enhanced food comprising one
 or more plant-derived food ingredients, and as an additive, a seed
 composition containing a flour, extract, or malt obtained from mature
 monocot seeds and one or more seed-produced human milk proteins in
 substantially unperfumed form. The nutritionally enhanced food and
 additive compositions are useful for improving infant formula. Infant
 formulas containing the recombinant human milk proteins are useful in
 supplementing or enhancing the diet of infants, particularly very-low
 birth weight infants. The present sequence is human codon optimised
 lactoferrin.

XX Sequence 692 AA;

Query Match

Best Local Similarity 97.8%; Score 261; DB 23; Length 692;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 48

|||||
 Db 3 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 49

RESULT 14

AAR85146

ID AAR85146 standard; Protein; 693 AA.

XX AAR85146;

XX 02-FEB-1996 (first entry)

XX Lactoferrin protein.

XX Lactoferrin; antiseptic; dietary supplement; iron sequestration;
 KW Pichia pastoris.

XX Homo sapiens.

XX WO9530339-A1.

XX 16-NOV-1995.

XX 05-MAY-1995; 95WO-US05653.

XX 05-MAY-1994; 94US-0238445.

XX (FERR-) FERRODYNAMICS INC.

XX Doyle D, Gollnick P, Kruzel ML, Kurecki T;

XX WPI; 1995-403881/51.

XX N-PSDB; AAT02999.

XX New isolated DNA encoding human lactoferrin protein - useful for
 inhibiting microbial growth and iron deficiency, and for
 sequestering iron or retarding food spoilage.

XX Disclosure; Fig 3a-d; 43pp; English.

XX cDNA encoding human lactoferrin (AAR85146) was isolated from a human
 mammary gland library and inserted into vector pUC118. Expression
 in Pichia pastoris GTS 115 (His4) allowed the prodn. of large
 amounts of the lactoferrin protein, useful as an antiseptic,
 dietary supplement, etc.

XX Sequence 693 AA;

Query Match

Best Local Similarity 97.8%; Score 261; DB 16; Length 693;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 48

|||||
 Db 4 RRRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 50

RESULT 15

AAR22423

ID AAR22423 standard; Protein; 705 AA.

XX AAR22423;

XX 25-MAR-2003 (updated)

XX 17-DEC-2001 (updated)

XX 05-AUG-1992 (first entry)

XX Human lactoferrin deduced from clone HLF 1212.

XX anti-bacterial; anti-viral; anti-fungal; Colony stimulating factor;
 KW CSF; release; leukaemia; breast cancer; hormonal regulation.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 22

XX /note= "Insertion, c.f. B.F.Anderson et al, 1989"
 XX Misc-difference 31

FT Misc-difference 55 /note= "Substitution, c.f. B.F.Anderson et al, 1989"
FT /note= "Substitution, c.f. B.F.Anderson et al.,1989"
FT Misc-difference 218 /note= "Substitution, c.f. B.F.Anderson et al.,1989"
FT /note= "Substitution, c.f. B.F.Anderson et al.,1989"

XX USN707502-N.
PN 18-FEB-1992.
XX 31-MAY-1991; 91US-0707502.
XX 31-MAY-1991; 91US-0707502.
PR (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX Teng C, Panella TJ;
XX WPT; 1992-113934/14.
DR N-PSDB; AAQ23453.
XX

XX Human lactoferrin gene and promoter - used for producing protein
PT for treating e.g. AIDS, neutropenia, skin infections, vaginal
PT infections or septic shock
XX Example 2; Fig 10; 55pp; English.

XX A Clonotech cDNA library from normal human breast tissue (HL1037b)
CC was plated in host cells 71090, filter-lifted and probed with mouse
CC lactoferrin cDNA T767. Positive clones were plaque-purified and the
CC inserts subcloned into the EcoRI site of Bluescript II SK+. The
CC recombinant clones were transformed into XLI Blue cells. The
CC longest insert (HLF 1212) was sequenced and was 2117bp in length.
CC The amino acid sequence coded for by HLF 1212 has 4 sites that
CC differ from the previously published revised amino acid sequence
CC derived from the protein (B.F.Anderson et al., J.Mol.Biol. 209:
CC 711-734, 1989). Also contained within HLF 1212, but not in any of
CC the other partially sequenced isolates, is a deleted cytosine at
CC bp 2097 (AA 699) which caused a frameshift at the 3' end of the coding
CC sequence. The deletion is thought to be either a cloning artefact
CC or a rare species of mRNA. See also AAQ23454 and AAQ23455.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 705 AA;

Query Match 97.8%; Score 261; DB 13; Length 705;
Best Local Similarity 100.0%; Pred. No. 3.8e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRSVQWCAVSQPEATKCFQWQRMKVRGPPVSCIKRDSPTQCTQA 48
Db ||||||||||||||||||||||||||||||||||||||||||||
20 RRSVQWCAVSQPEATKCFQWQRMKVRGPPVSCIKRDSPTQCTQA 66

Search completed: August 11, 2003, 08:18:24
Job time : 36.9515 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:16:37 ; Search time 19,1068 Seconds
(without alignments)
106.293 Million cell updates/sec

Title: US-09-508-095-19
Perfect score: 267
Sequence: 1 GRRSVQWCAVSQPEATKCF.....VRGPPVSCIKRDSPIQIQIA 48

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	97.8	54	2	US-08-464-182A-2
2	261	97.8	54	2	US-08-406-271-2
3	261	97.8	694	3	US-08-724-586-2
4	261	97.8	694	3	US-08-421-632-2
5	261	97.8	694	3	US-08-932-190-2
6	261	97.8	705	2	US-08-655-840-2
7	261	97.8	708	2	US-08-655-840-4
8	261	97.8	711	1	US-08-154-019-4
9	261	97.8	711	1	US-08-461-333-4
10	261	97.8	711	3	US-08-464-167-4
11	261	97.8	711	3	US-08-158-313-4
12	261	97.8	711	3	US-08-476-798-4
13	260	97.4	53	2	US-08-464-182A-5
14	260	97.4	53	2	US-08-406-271-5
15	254	95.1	52	4	US-08-017-043A-3
16	254	95.1	709	1	US-08-154-019-2
17	254	95.1	709	1	US-08-461-333-2
18	254	95.1	709	3	US-08-464-167-2
19	254	95.1	709	3	US-08-158-313-2
20	254	95.1	709	3	US-08-476-798-2
21	254	95.1	711	1	US-08-145-681-2
22	254	95.1	711	1	US-08-250-308-2
23	254	95.1	711	1	US-08-453-703-2
24	254	95.1	711	3	US-08-456-106-2
25	254	95.1	711	3	US-08-456-108-2
26	254	95.1	711	3	US-08-265-577-2
27	254	95.1	711	5	PCT-US93-03614-2

28	252	94.4	47	2	US-08-464-182A-6	Sequence 6, Appli
29	252	94.4	47	2	US-08-406-271-6	Sequence 6, Appli
30	250	93.6	30	2	US-08-693-274A-7	Sequence 7, Appli
31	200	74.9	36	1	US-07-755-161A-8	Sequence 8, Appli
32	200	74.9	36	1	US-07-891-174-8	Sequence 8, Appli
33	200	74.9	36	1	US-08-256-771-30	Sequence 30, Appli
34	200	74.9	36	1	US-08-381-984-29	Sequence 29, Appli
35	159	59.6	48	2	US-08-693-274A-10	Sequence 10, Appli
36	159	59.6	50	4	US-09-017-043A-5	Sequence 5, Appli
37	159	59.6	52	2	US-08-464-182A-17	Sequence 17, Appli
38	159	59.6	52	2	US-08-406-271-17	Sequence 17, Appli
39	142	53.2	25	1	US-07-755-161A-10	Sequence 10, Appli
40	142	53.2	25	1	US-07-891-174-10	Sequence 10, Appli
41	142	53.2	25	1	US-08-204-487-7	Sequence 7, Appli
42	141	52.8	30	2	US-08-464-182A-4	Sequence 4, Appli
43	141	52.8	30	2	US-08-406-271-4	Sequence 4, Appli
44	134	50.2	51	4	US-09-017-043A-4	Sequence 4, Appli
45	134	50.2	53	2	US-08-464-182A-22	Sequence 22, Appli

ALIGNMENTS

RESULT 1
US-08-464-182A-2
; Sequence 2, Application US/08464182A
; Patent No. 5849885
; GENERAL INFORMATION:
; APPLICANT: Nuljens, Jan H.
; APPLICANT: Van Veen, Harry H.
; TITLE OF INVENTION: Isolation of Lactoferrin from Milk
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,182A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/406,271
; FILING DATE: 09-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,321
; FILING DATE: 16-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 016994-0049200S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..54
; OTHER INFORMATION: /note= "HLF cdNA of GPE"

US-08-464-182A-2

Query Match 97.8%; Score 261; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 8e-29;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 48
Db 3 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 49

RESULT 2

US-08-406-271-2
; Sequence 2, Application US/08406271
; Patent No. 5919913

GENERAL INFORMATION:
APPLICANT: Nuyens, Jan H.

APPLICANT: Van Veen, Harry H.
TITLE OF INVENTION: Isolation of Lactoferrin from Milk

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: One Market Plaza, Steuart Street Tower, 20th

CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/406,271
FILING DATE: 09-MAR-1995

CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/198,321
FILING DATE: 16-FEB-1994

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 016994-004910

TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 54 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant

MOLECULE TYPE: peptide
FEATURE:

NAME/KEY: Peptide
LOCATION: 1..54

OTHER INFORMATION: /note= "HLF cDNA of GPE"
US-08-406-271-2

Query Match 97.8%; Score 261; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 8e-29;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 48
Db 3 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 49

RESULT 3

US-08-724-586-2

; Sequence 2, Application US/08724586
; Patent No. 6066469

GENERAL INFORMATION:

APPLICANT: Kruzel, Marian L.

APPLICANT: Kurecki, Tomasz

APPLICANT: Gollnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/724,586

FILING DATE: 30-SEPT-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/238,445

FILING DATE: 05-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,409

REFERENCE/DOCKET NUMBER: 10505/P58185C

TELEPHONE: (202) 638-6666

TELEFAX: (202) 393-5350

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-724-586-2

Query Match 97.8%; Score 261; DB 3; Length 694;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 48
Db 5 RRRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQICQA 51

RESULT 4

US-09-421-632-2

; Sequence 2, Application US/09421632
; Patent No. 6277817

GENERAL INFORMATION:
APPLICANT: Kruzel, Marian L.

APPLICANT: Kurecki, Tomasz

APPLICANT: Gollnick, Paul D.

APPLICANT: Doyle, Darrell J.

TITLE OF INVENTION: Cloning, Expression, and Uses of Human

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman & Stern

STREET: 400 Seventh St. N.W.

CITY: Washington D.C.

COUNTRY: U.S.A.

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/09/421,632
, FILING DATE:
, CLASSIFICATION:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/724,586
, FILING DATE:
, ATTORNEY/AGENT INFORMATION:
, NAME: Player, William E.
, REGISTRATION NUMBER: 31,409
, REFERENCE/DOCKET NUMBER: 10505/P51
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (202) 638-6666
, TELEFAX: (202) 393-5350
, INFORMATION FOR SEQ ID NO: 2:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 694 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: protein
US-09-421-632-2

```

```
Query Match          97.8%; Score 261; DB 3; Length 694;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 47; Conservative 0; Mismatches 0; Indels

QY 2 RRRSVQCVCVSQPEATKCFQWRNNRKVRGPPVSCIKRDSPICQIQQA 48
      |||||
Db    5 RRRSVQCVCVSQPEATKCFQWRNNRKVRGPPVSCIKRDSPICQIQQA 51
```

RESULT 5.
 US-09-932-190-2
 ; Sequence 2, Application US/09932190
 ; Patent No. 6455687
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruzel, Marian L.
 ; Kurecki, Tomasz
 ; Gollnick, Paul D.
 ; Doyle, Darrell J.
 ; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
 ; Lactoferrin
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Jacobson, Price, Holman & Stern
 ; STREET: 400 Seventh St. N.W.
 ; CITY: Washington D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/932.190
 ; FILING DATE: 17-Aug-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/724.586
 ; FILING DATE: 30-SEPT-1996
 ; APPLICATION NUMBER: US 08/238,445
 ; FILING DATE: 05-MAY-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Player, William E.
 ; REGISTRATION NUMBER: 31,409
 ; REFERENCE/DOCKET NUMBER: 10505/P58185C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 638-6666
 ; TELEFAX: (202) 393-3350
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 694 amino acids

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-932-190-2

Query Match          97.8%; Score 261; DB 4; Length 694;
Best Local Similarity 100.0%; Pred. No. 1.4e-27;
Matches 47; Conservative 0; Mismatches 0; Indels

QY      2  RRRSVQNCVQSQPATKCFQQRNNRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db      5  RRRSVQNCVQSQPATKCFQQRNNRKVRGPPVSCIKRDSPIQCIQA 51
|||||

```

RESULT 6
US-08-655-640-2
; Sequence 2, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,640
; FILING DATE: 30-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,538
; FILING DATE: December 17, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-655-640-2

```

Query Match          97.8%; Score 261; DB 2; Length 705;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels

QY      2   RRSVQMCVSPQEPATKCFQWRNMNRKVRGPPVSCIKRDSPIQCIQA 48
        |||||
Db       20   RRSVQMCVSPQEPATKCFQWRNMNRKVRGPPVSCIKRDSPIQCIQA 66

```

APPLICANT: Teng, Christina
APPLICANT: Panella, Timothy J.
TITLE OF INVENTION: HUMAN LACTOFERRIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,640
FILING DATE: 30-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/992,538
FILING DATE: December 17, 1992
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 708 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-640-4

Query Match 97.8%; Score 261; DB 2; Length 708;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQCIQA 48
Db 22 RRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQCIQA 68
|||||

RESULT 8
US-08-154-019-4
Sequence 4, Application US/08154019
Patent No. 5633076
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
TITLE OF INVENTION: by Bovine Species and Transgenic Methods
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,019
FILING DATE: 16-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/461,333
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebescheutz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-154-019-4

Query Match 97.8%; Score 261; DB 1; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQCIQA 48
Db 22 RRSVQWCAVSQPEATKCFQWQNRNMRKVRGPPVSCIKRDSPIQCIQA 68
|||||

RESULT 9
US-08-461-333-4
Sequence 4, Application US/08461333
Patent No. 5741957
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
TITLE OF INVENTION: by Bovine Species and Transgenic Methods
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,333

; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,788
; FILING DATE: 15-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/895,956
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,745
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebescheutz, Joe O.
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16994-003123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-333-4

Query Match 97.8%; Score 261; DB 1; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRRSVQWCAVSQPEATKCFQWQNRNKRKVRGPPVSCIKRDSPIQCIQA 48
DB 22 RRRSVQWCAVSQPEATKCFQWQNRNKRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 10
US-08-464-167-4
; Sequence 4, Application US/08464167
; Patent No. 6013857
; GENERAL INFORMATION:
; APPLICANT: Deboer, Herman A.
; APPLICANT: Strijker, Rein
; APPLICANT: Heyneker, Herbert L.
; APPLICANT: Platenburg, Gerald
; APPLICANT: Lee, Sang He
; APPLICANT: Pieper, Frank
; APPLICANT: Krimpenfort, Paul J.A.
; TITLE OF INVENTION: Production of Recombinant Polypeptides
; TITLE OF INVENTION: by Bovine Species and Transgenic Methods
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,167
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,788
; FILING DATE: 15-JUN-1993
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/895,956
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,745
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebescheutz, Joe O.
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16994-003124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-167-4

Query Match 97.8%; Score 261; DB 3; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRRSVQWCAVSQPEATKCFQWQNRNKRKVRGPPVSCIKRDSPIQCIQA 48
DB 22 RRRSVQWCAVSQPEATKCFQWQNRNKRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 11
US-09-158-313-4
; Sequence 4, Application US/09158313
; Patent No. 6066725
; GENERAL INFORMATION:
; APPLICANT: Deboer, Herman A.
; APPLICANT: Strijker, Rein
; APPLICANT: Heyneker, Herbert L.
; APPLICANT: Platenburg, Gerald
; APPLICANT: Lee, Sang He
; APPLICANT: Pieper, Frank
; APPLICANT: Krimpenfort, Paul J.A.
; TITLE OF INVENTION: Production of Recombinant Polypeptides
; TITLE OF INVENTION: by Bovine Species and Transgenic Methods
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/158,313
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/895,956
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschultz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003125
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-158-313-4

Query Match 97.8%; Score 261; DB 3; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 48
DB 22 RRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 12
US-08-476-798-4
Sequence 4, Application US/08476798
Patent No. 6140552
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,798
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschultz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003125

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-798-4

Query Match 97.8%; Score 261; DB 3; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.5e-27;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 48
DB 22 RRSVQWCAVSQPEATKCFQWRNMRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 13
US-08-464-182A-5
Sequence 5, Application US/08464182A
Patent No. 5849885
GENERAL INFORMATION:
APPLICANT: Nuijens, Jan H.
APPLICANT: Van Veen, Harry H.
TITLE OF INVENTION: Isolation of Lactoferrin from Milk
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,182A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/406,271
FILING DATE: 09-MAR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,321
FILING DATE: 16-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 016994-0049200S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 326-2400
TELEFAX: (650) 326-2422
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1-53
OTHER INFORMATION: /note= "HLF CDNA Rado"
US-08-464-182A-5

Query Match 97.4%; Score 260; DB 2; Length 53;

Best Local Similarity 95.8%; Pred. No. 1.1e-28;
Matches 46; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRRSVQWCAVSOPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 1 GRRSVQWCAVSNPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 48

RESULT 14

US-08-406-271-5
; Sequence 5, Application US/08406271
; Patent No. 5919913
; GENERAL INFORMATION:
; APPLICANT: Nuijens, Jan H.
; APPLICANT: Van Veen, Harry H.
; TITLE OF INVENTION: Isolation of Lactoferrin from Milk
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: One Market Plaza, Steuart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,271
; FILING DATE: 09-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,321
; FILING DATE: 16-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 016994-004910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..53
; OTHER INFORMATION: /note= "HLF cdna Rado"
US-08-406-271-5

Query Match 97.4%; Score 260; DB 2; Length 53;
Best Local Similarity 95.8%; Pred. No. 1.1e-28;
Matches 46; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRRSVQWCAVSOPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 1 GRRSVQWCAVSNPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 48

RESULT 15

US-09-017-043A-3
; Sequence 3, Application US/09017043A
; Patent No. 6333311
; GENERAL INFORMATION:
; APPLICANT: Nuijens, Jan
; APPLICANT: van Berkel, Patrick

; TITLE OF INVENTION: Useful Properties of Human Lactoferrin
; TITLE OF INVENTION: and Variants Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,043A
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,859
; FILING DATE: 03-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 016994-011010US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-017-043A-3

Query Match 95.1%; Score 254; DB 4; Length 52;
Best Local Similarity 95.7%; Pred. No. 6.9e-28;
Matches 45; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSOPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 3 RRRSVQWCAVSNPEATKCFQWQRMNRKVRGPPVSCIKRDSPIQCIQA 49

Search completed: August 11, 2003, 08:19:45
Job time : 20.1068 secs

16	76	28.5	705	11	US-09-847-208-102	Sequence 102, Appl
17	70	26.2	328	9	US-09-891-126-5	Sequence 5, Appl1
18	65	24.3	23	9	US-09-809-652-4	Sequence 4, Appl1
19	65	24.3	25	9	US-030-619-205	Sequence 205, App
20	65	24.3	25	9	US-09-917-340-16	Sequence 16, Appl
21	65	24.3	25	11	US-09-798-869-21	Sequence 21, Appl
22	60	22.5	15	11	US-09-798-869-3	Sequence 3, Appl1
23	54	20.2	15	11	US-09-798-869-4	Sequence 4, Appl1
24	54	20.2	49	9	US-09-864-761-45442	Sequence 45442, A
25	53	19.9	48	10	US-09-941-314-17	Sequence 17, Appl
26	53	19.9	59	10	US-09-941-314-16	Sequence 16, Appl
27	53	19.9	80	10	US-09-941-314-15	Sequence 15, Appl
28	53	19.9	115	10	US-09-941-314-3	Sequence 3, Appl1
29	53	19.9	117	10	US-09-941-314-4	Sequence 4, Appl1
30	53	19.9	137	10	US-09-941-314-2	Sequence 2, Appl1
31	53	19.9	188	9	US-09-925-301-1331	Sequence 1331, Ap
32	53	19.9	258	10	US-09-731-872-313	Sequence 313, App
33	53	19.9	337	10	US-09-731-872-248	Sequence 248, App
34	52.5	19.7	64	9	US-09-917-340-87	Sequence 87, Appl
35	52	19.5	15	11	US-09-798-869-7	Sequence 7, Appl1
36	52	19.5	103	11	US-09-975-719-291	Sequence 291, App
37	51	19.1	15	11	US-09-798-869-8	Sequence 8, Appl1
38	51	19.1	128	9	US-09-864-761-36308	Sequence 36308, A
39	51	19.1	284	16	US-10-080-170-388	Sequence 388, App
40	50	18.7	132	15	US-10-311-111-21	Sequence 21, Appl
41	50	18.7	306	15	US-10-156-761-8748	Sequence 8748, Ap
42	50	18.7	393	10	US-09-764-868-690	Sequence 690, App
43	49.5	18.5	64	9	US-09-917-340-88	Sequence 88, Appl
44	48.5	18.2	173	15	US-10-097-065-252	Sequence 252, App
45	48.5	18.2	173	15	US-10-097-065-525	Sequence 525, App

```

RESULT 1
US-10-076-816-4
; Sequence 4, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagite, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076.816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,198
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
SOFTWARE: FASTSEQ for Windows Version 4.0

```

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	261	97.8	690	US-10-076-816-4	Sequence 4, Appli
2	261	97.8	690	US-10-077-381-4	Sequence 4, Appli
3	261	97.8	694	US-10-023-096-2	Sequence 2, Appli
4	142	53.2	25	US-09-788-869-20	Sequence 20, Appli
5	118.5	44.4	164	US-09-738-973-203	Sequence 203, App
6	118.5	44.4	164	US-09-854-133-203	Sequence 203, App
7	118.5	44.4	164	US-10-144-649A-203	Sequence 203, App
8	118.5	44.4	698	US-09-935-642-6	Sequence 6, Appli
9	106.5	39.9	255	US-09-866-050A-693	Sequence 693, App
10	96	36.0	25	US-09-798-869-22	Sequence 22, Appl
11	90	33.7	25	US-09-798-869-23	Sequence 23, Appl
12	87	32.6	15	US-09-798-869-2	Sequence 2, Appli
13	86	32.2	17	US-10-145-651-1	Sequence 1, Appli
14	79	28.6	15	US-09-798-869-6	Sequence 6, Appli
15	76.5	28.7	719	US-10-193-642-3	Sequence 3, Appli

US-10-077-381-4
; Sequence 4, Application US/10077381
; Publication No. US20030074700A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Expression of Human Milk Proteins in
; Transgenic Plants
; FILE REFERENCE: 50665-8022 US00
; CURRENT APPLICATION NUMBER: US/10/077,381
; CURRENT FILING DATE: 2002-02-14
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2001-05-02
; PRIOR FILING DATE: 2001-02-06
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-381-4

Query Match 97.8%; Score 261; DB 15; Length 690;
Best Local Similarity 100.0%; Pred. No. 5.9e-26;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 48
|||||
DB 3 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 49
|||||

RESULT 3
US-10-023-096-2
; Sequence 2, Application US/10023096
; Publication No. US20020160941A1
; GENERAL INFORMATION:
; APPLICANT: Kruzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Gollnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,096
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: 10505/P58185C

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-023-096-2

Query Match 97.8%; Score 261; DB 14; Length 694;
Best Local Similarity 100.0%; Pred. No. 5.9e-26;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 48
|||||
DB 5 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 51
|||||

RESULT 4
US-09-798-869-20
; Sequence 20, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (ISTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GH99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-798-869-20

Query Match 53.2%; Score 142; DB 11; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TKCFQWQNRMRKVRGPPVSCIKRDS 41
|||||
DB 1 TKCFQWQNRMRKVRGPPVSCIKRDS 25
|||||

RESULT 5
US-09-738-973-203
; Sequence 203, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR LUNG CANCER
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF

Matches	20;	Conservative	13;	Mismatches	12;	Indels	3;	Gaps	1;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

Qy 4 RSYQWCAVSQPEATKCFQWQNRKV---RGPPVSVCKRDSPTQCTQA 48
:::|||||:| ||| :::::| || :::::| |||
Db 23 KTKVCAVSEHENTKICISFDHMKTVLPDGGPLACVKKTSYDQIKA 70

RESULT 10

```

US-09-798-869-22
; Sequence 22, Application US/09798869
; Publication No. US2003002821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJARNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/G899/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: G89818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 25
; TYPE: PRT
; ORGANISM: MURINE
US-09-798-869-22

```

Query Match 36.0%; Score 95; DB 11; Length 25;
 . Best Local Similarity 62.5%; Pred. No. 7.8e-06;
 Matches 15; Conservative 4; Mismatches 5; Indels

Qy 18 KCFWQNRMRKVRGPPVSCIKRDS 41
|| :|| |||| ||||:|:|
Db 2 KCLRWQNRMRKVGPPPLSCVKKSS 25

RESULT 11

```

US-09-798-869-23
; Sequence 23, Application US/09798869
; Publication No. US2003002821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNNSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 25
; TYPE: PRT
; ORGANISM: CAPRINE
US-09-798-869-23

```

Query Match 33.7%; Score 90; DB 11; Length 25;
Best Local Similarity 52.0%;
Matches 13; Conservative 7; Mismatches 5; Indels
Pred. No. 4.8e-05;

```
.      QY          17 TKCFQWQRNMRKVRGPPVSCIKRDS   41  
           ::::| | | | | ::|||  
Db       1 SKCYQWQRMRMKLGAPSI TCVRRTS    25
```

RESULT 12

```

US-09-798-869-2
; Sequence 2, Application US/0979869
; Publication No. US2003002821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJORNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9819938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-798-869-2

```

Query Match 32.6%; Score 87; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels

QY 17 TKCFQWQRNMRKVRG 31
Db 1 TKCFQWQRNMRKVRG 15

RESULT 13

```

US-10-145-651-1
; Sequence 1, Application US/10145651
; Publication NO. US20030105006A1
; GENERAL INFORMATION:
; APPLICANT: Mann, David
; TITLE OF INVENTION: ANTIMICROBIAL/ENDOTOXIN NEUTRALIZING
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: 00138.0051.NPUS00
; CURRENT APPLICATION NUMBER: US/10/145,651
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US/09/245,527
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-651-1

```

Query Match	32.2%	Score 86;	DB 15;	Length 17;
Best Local Similarity	100.0%;	Pred. No. 0.00011;		
Matches	16; Conservative	0; Mismatches	0; Indels	

Qy	33	PVSCIKRDSPIQCIQA	48
Db	1	PVSCIKRDSPIQCIQA	16

RESULT 14

US-09-798-869-6
; Sequence 6, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: ØYSTEIN REXDAL

Job time : 132.35 secs

```
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens
; OTHER INFORMATION: sequence)
US-09-798-869-6
```

```
Query Match 29.6%; Score 79; DB 11; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.00078;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 TKCFQWQNRMRKVRG 31
Db 1 TKCFQWQNRMRKVRG 15
```

```
RESULT 15
US-10-133-642-3
; Sequence 3, Application US/10133642
; Publication No. US20030077693A1
; GENERAL INFORMATION:
; APPLICANT: Jefferies, Wilfred A.
; APPLICANT: McGeer, Patrick L
; APPLICANT: Rothenberger, Sylvia
; APPLICANT: Food, Michael R.
; APPLICANT: Yamada, Tatsuo
; APPLICANT: Kennard, Malcolm
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins as Diagnostic and Therapeutic
; FILE REFERENCE: 7685-48
; CURRENT APPLICATION NUMBER: US/10/133,642
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/285,040
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: US 08/520,933
; PRIOR FILING DATE: 1995-08-31
; PRIOR APPLICATION NUMBER: US 08/367,224
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/CA93/00272
; PRIOR FILING DATE: 1995-03-30
; PRIOR APPLICATION NUMBER: US 07/912,291
; PRIOR FILING DATE: 1992-07-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-642-3
```

```
Query Match 28.7%; Score 76.5; DB 15; Length 719;
Best Local Similarity 32.6%; Pred. No. 0.091;
Matches 14; Conservative 8; Mismatches 20; Indels 1; Gaps 1;

Qy 6 VQWCAVSQPEATKCFQWQNRMRKVR-GPPVSCIKRDSPIQIQ 47
Db 347 LRWCVLSTPEIQKCGDMAVAFRRQRRLKPEIQCVSAKSPQHME 389
```

Search completed: August 11, 2003, 08:25:41

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:16:37 ; Search time 13.0485 Seconds
(without alignments)
353.763 Million cell updates/sec

Title: US-09-508-095-19

Perfect score: 267

Sequence: 1 GRRRSVQWCAVSQPEATKCF.....VRGPPVSCIKRDSPIQCIQA 48

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	97.8	711	1 TFHUL	lactotransferrin p
2	162	60.7	708	2 JC2323	lactoferrin - goat
3	159	59.6	707	1 A28438	lactoferrin p
4	134	50.2	708	1 TFHUL	lactotransferrin p
5	132.5	49.6	704	2 I47228	carbonic anhydrase
6	120	44.9	696	1 S01384	transferrin - pig
7	118.5	44.4	698	1 TFHUP	transferrin p
8	118	44.2	703	2 A45543	lactoferrin p
9	115	43.1	33	2 S52107	lactoferrin p
10	110.5	41.4	694	1 TFHUP	transferrin - sheep
11	104.5	39.1	695	2 S49163	transferrin p
12	103	38.6	706	2 S33761	transferrin p
13	76.5	28.7	64	2 I51350	transferrin - Atla
14	76.5	28.7	690	2 T11749	transferrin - Atla
15	76.5	28.7	738	1 TFHUM	melanotransferrin
16	76	28.5	705	1 TFCHE	ovotransferrin p
17	72	27.0	717	2 S12100	transferrin p
18	68	25.5	311	2 A28446	transferrin - mous
19	60	22.5	551	2 G82630	fumarate hydratase
20	57	21.3	373	1 S29934	chorismate mutase
21	57	21.3	373	2 A10832	prephenate dehydro
22	57	21.3	346	2 T19139	hypothetical prote
23	56	21.0	262	2 C75554	conserved hypothet
24	55	20.6	2505	1 XYRTPA	enoyl-[acyl-carrie
25	54	20.2	417	2 T24321	hypothetical prote
26	54	20.2	419	2 T24322	hypothetical prote
27	54	20.2	422	2 A34569	lin-10 protein - C
28	53.5	20.0	1371	1 VCBEM7	major capsid prote
29	53.5	20.0	1371	2 T42938	major capsid prote

30 53.5 20.0 3871 2 T22812
31 53 19.9 170 2 AB2715
32 53 19.9 189 2 F97496
33 53 19.9 370 1 LNRBB
34 53 19.9 373 1 KMECTD
35 53 19.9 373 2 C85906
36 53 19.9 373 2 G91081
37 52.5 19.7 388 2 T29694
38 52.5 19.7 738 2 T49295
39 52.5 19.7 820 2 A86247
40 51.5 19.3 681 1 A36500
41 51.5 19.3 811 2 T40275
42 51.5 19.3 815 2 G82417
43 51.5 19.3 1265 2 A59299
44 51 19.1 274 2 JQ2159
45 51 19.1 274 2 JQ2145

ALIGNMENTS

RESULT 1

TFHUL
lactotransferrin precursor [validated] - human
N:Alternate names: lactoferrin
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 21-Nov-1997 #text_change 08-Dec-2000
C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000;
R:Cho, Y.
submitted to the EMBL Data Library, March 1994
A:Reference number: G06820
A:Accession: G01394
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-711 <CHO>
A:Cross-references: EMBL:U07643; NID:9467236; PIDN:AAB60324.1; PID:9467237
R:Ray, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.
Nucleic Acids Res. 18, 5288, 1990
A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.
A:Reference number: S11228; MUID:90384839; PMID:2402455
A:Accession: S11228
A:Molecule type: mRNA
A:Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>
A:Cross-references: EMBL:X53961; NID:934415; PIDN:CAA37914.1; PID:934416
R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.
Mol. Endocrinol. 6, 1969-1981, 1992
A:Title: Differential molecular mechanism of the estrogen action that regulates lacto
A:Reference number: A45401; MUID:93125571; PMID:1480183
A:Accession: A45401
A:Molecule type: DNA
A:Residues: 1-15 <TEN>
A:Cross-references: GB:S52659; NID:9263311; PIDN:AAB24877.1; PID:9263312
A:Experimental source: placenta
A:Note: sequence extracted from NCBI backbone (NCBIP:122202)
R:Powell, M.J.; Ogden, J.E.
Nucleic Acids Res. 18, 4013, 1990
A:Title: Nucleotide sequence of human lactoferrin cDNA.
A:Reference number: S10324; MUID:90326549; PMID:2374734
A:Accession: S10324
A:Molecule type: mRNA
A:Residues: 3-711 <POW>
A:Cross-references: EMBL:X52941; NID:934411; PIDN:CAA37116.1; PID:934412
R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.
Biochem. J. 276, 349-355, 1991
A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.
A:Reference number: S15853; MUID:91264786; PMID:2049066
A:Accession: S15853
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 20-31 <ST1>
A:Accession: S20841
A:Molecule type: protein
A:Residues: 20-28, 'X', 30-31 <ST2>

```

R:Rado, T.A.; Wei, X.; Benz Jr., E.J.
Blood 70, 989-993, 1987
A:Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA
A:Reference number: S07160; MUID:88001031; PMID:3477300
A:Accession: S07160
A:Molecule type: mRNA
A:Residues: 436-487, 'A', 489-711 <RAD>
A:Cross-references: EMBL:M18642; NID:g186815; PIDN:AAA86665.1; PID:g386855
R:Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.
Cancer Res. 51, 3037-3043, 1991
A:Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes
A:Reference number: A61169; MUID:91235214; PMID:1674448
A:Accession: A61169
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 3-701, 'SWKPVN' <PAN>
R:Experimental source: normal breast tissue
R:Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;
Eur. J. Biochem. 145, 659-666, 1994
A:Title: Human lactotransferrin: amino acid sequence and structural comparisons with other
A:Reference number: A31000; MUID:85076667; PMID:6510420
A:Accession: A31000
A:Molecule type: protein
A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4
A:Note: this is the final paper in a series
R:Houen, G.; Hoegdall, E.V.; Barkholt, V.; Nørskov, L.
Eur. J. Biochem. 241, 303-308, 1996
A:Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity
A:Reference number: S74119; MUID:97054624; PMID:8898921
A:Accession: S74119
A:Molecule type: protein
A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>
A:Experimental source: neutrophil granulocytes
C:Genetics:
A:Gene: GDB:LNF
A:Cross-references: GDB:110368; OMIM:150210
A:Map position: 3q21-3q23
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein; iron binding; milk
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-711/Product: lactotransferrin #status experimental <MAT>
F:21-356/Domain: transferrin repeat homology <TRH2>
F:360-699/Domain: transferrin repeat homology <TRH2>
F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status
F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

Query Match 97.8%; Score 261; DB 1; Length 711;
Best Local Similarity 100.0%; Pred. No. 3.4e-25;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 22 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 2
JC2323
lactoferrin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C:Accession: JC2323
R:Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.
Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994
A:Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus
A:Reference number: JC2323; MUID:94380047; PMID:8093048
A:Accession: JC2323
A:Molecule type: mRNA
A:Residues: 1-708 <LEP>
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein
F:359-696/Domain: transferrin repeat homology <TRH2>
F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.7%; Score 162; DB 2; Length 708;
Best Local Similarity 52.2%; Pred. No. 1.2e-12;
Matches 24; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 3 RRSVQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 22 RKNVRWCAISLPENSKCYQWQNRMRKLGAPSTICVRRTSALECIRA 67

RESULT 3
A28438
lactoferrin precursor - mouse
N:Alternate names: lactotransferrin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A28438; A41205
R:Pentecost, B.T.; Teng, C.T.
J. Biol. Chem. 262, 10134-10139, 1987
A:Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine se
A:Reference number: A92596; MUID:87280033; PMID:3611056
A:Accession: A28438
A:Molecule type: mRNA
A:Residues: 3-707 <PEN>
A:Cross-references: EMBL:J03298
R:Liu, Y.; Teng, C.T.
J. Biol. Chem. 266, 21880-21885, 1991
A:Title: Characterization of estrogen-responsive mouse lactoferrin promoter.
A:Reference number: A41205; MUID:92042099; PMID:1939212
A:Accession: A41205
A:Molecule type: DNA
A:Residues: 1-15 <LIU>
A:Cross-references: GB:M74778
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-707/Product: lactotransferrin #status predicted <MAT>
F:358-695/Domain: transferrin repeat homology <TRH2>
F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 59.6%; Score 159; DB 1; Length 707;
Best Local Similarity 61.4%; Pred. No. 2.8e-12;
Matches 27; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 5 SYQWCAVSQPEATKCFQWQNRMRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 23 TYRWCAVNSEEEKLRQWQNRMRKVRGPPVSCVKKSTRQCIQA 66

RESULT 4
TFBOL
lactotransferrin precursor - bovine
N:Alternate names: lactoferrin
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Mar-1992 #sequence_revision 21-Nov-1997 #text_change 11-May-2000
C:Accession: I45919; S14674; S18517; J0595; S13097; S18518; S13881; PL0148;
R:Tsang, T.C.; Burns, D.K.; Wang, F.; Pan, Y.
FASEB J. 6, 233, 1991
A:Title: Cloning of a 80-kD advanced glycosylation end product (AGE) binding protein
A:Reference number: I45919
A:Accession: I45919
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-708 <TSA>
A:Cross-references: GB:L08604; NID:g163269; PIDN:AAA30609.1; PID:g163270
R:Pierce, A.
submitted to the EMBL Data Library, November 1990
A:Reference number: S14674
A:Accession: S14674
A:Molecule type: mRNA
A:Residues: 1-144, 'V', 146-163, 'PP', 166-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI>
A:Cross-references: EMBL:X57084; NID:g505; PIDN:CAA40366.1; PID:g506
R:Pierce, A.; Colavizza, D.; Benaissa, M.; Maes, P.; Tartar, A.; Montreuil, J.; Spik,

```

RESULT 6

S01384

transferrin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text_change 11-May-2000

C:Accession: S01384; A60520; A61573

R:Baldwin, G.S.; Weinstein, J.

Nucleic Acids Res. 16, 8720, 1988

A:Title: Nucleotide sequence of porcine liver transferrin.

A:Reference number: S01384; MUID:88335629; PMID:3419934

A:Accession: S01384

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-696 <BAL>

A:Cross-references: EMBL:X12386; NID:g2126; PIDN:CAA30943.1; PID:g833800

A:Note: 308-Arg was also found

R:Baldwin, G.S.; Baclic, T.; Chandler, R.; Grego, B.; Pedersen, J.; Simpson, R.J.; Toghiani, H.

Comp. Biochem. Physiol. B 95, 261-268, 1990

A:Title: Isolation of transferrin from porcine gastric mucosa: comparison with porcine transferrin.

A:Reference number: A60520; MUID:90227903; PMID:2328566

A:Accession: A60520

A:Molecule type: protein

A:Residues: 1-8,'X',10-11,'X',13-15 <BA2>

A:Experimental source: gastric mucosa

A:Note: the authors suggest transferrin from gastric mucosa may act in dietary iron uptake.

R:Chung, M.C.M.; Chan, S.L.; Shimizu, S.

Int. J. Biochem. 23, 609-616, 1991

A:Title: Purification of transferrins and lactoferrin using DEAE Affi-Gel Blue.

A:Reference number: A61573; MUID:91293379; PMID:2065820

A:Accession: A61573

A:Molecule type: protein

A:Residues: 1-8, 'X',10-18, 'XE' <CHU>

C:Superfamily: transferrin; transferrin repeat homology

C:Keywords: duplication; glycoprotein; iron transport; plasma

R;Yoshinari, K.; Yuasa, K.; Iga, F.; Mimura, A.
Biochim. Biophys. Acta 1010, 28-34, 1989
A;Title: A growth-promoting factor for human myeloid leukemia cells from horse serum
A;Reference number: S02145; MUID:89076897; PMID:2909248
A;Accession: S02145
A:Molecule type: protein
A;Residues: 20-35,'X','33','X',39-40,'X',43-44 <YOS>
C;Complex: monomer
C;Superfamily: transferrin; transferrin repeat homology
C;Keywords: duplication; glycoprotein; iron transport; metal binding; plasma
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-706/Product: transferrin #status experimental <MAT>
F;358-694/Domain: transferrin repeat homology <TRH>
F;26-64,36-55,134-215,174-190,177-198,187-200,248-262,360-623,366-398,376-389,423-703

Query Match 38.6%; Score 103; DB 2; Length 706;
Best Local Similarity 42.6%; Pred. No. 3.4e-05;
Matches 20; Conservative 13; Mismatches 12; Indels 2; Gaps 2;

QY 4 RSVQWCAVSQPATKCFQWQRNMRK-VRGPP-VSCIARDSPICQIA 48
 :::||| | : : : : : : : | | : : || : : || : :
Db 21 QTVRMCTVSNHEVSKCAFSDRMKSIVPAPPLVACVRKTSYLECIKA 67

RESULT 13
I51350
transferrin - Atlantic salmon (fragment)
C;Species: Salmo salar (Atlantic salmon)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51350
R;Kvingedal, A.M.
Gene 150, 335-339, 1994
A;Title: Characterization of the 5' region of the Atlantic salmon (Salmo salar) trans-
A;Reference number: I51350; MUID:95121925; PMID:7821802
A;Accession: I51350
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-64 <KVI>
A;Cross-references: GB:L26909; NID:g598395; PIDN:AAC42221.1; PID:g598396
C;Genetics:
A;Gene: Tf
A;Introns: 14/1
C;Superfamily: transferrin; transferrin repeat homology

Query Match 28.7%; Score 76.5; DB 2; Length 64;
Best Local Similarity 32.6%; Pred. No. 0.0085;
Matches 14; Conservative 8; Mismatches 16; Indels 5; Gaps 1;

QY 6 VQWCAVSQPATKCFQWQRNMRKVRGPPVSCIRDSPIQCIOA 48
 | : | | : | : | : : : : : : : : : : : : : :
Db 25 VKMKVSEQLRKCHDLAAKVAE-----FSCVRKDGSFEQIOA 62

RESULT 14
T11749
transferrin - Atlantic salmon
C;Species: Salmo salar (Atlantic salmon)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T11749
Mol. Marine Biol. Biotechnol. 2, 233-238, 1993
R;Kvingedal, A.M.; Rorvik, K.A.; Alestrom, P.
A;Title: Cloning and characterization of Atlantic salmon (Salmo salar) serum transferri-
A;Reference number: Z17332; MUID:94122797; PMID:8293074
A;Accession: T11749
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-690 <KVI>
A;Cross-references: EMBL:L20313; NID:g431609; PIDN:AAA18838.1; PID:g431610
A;Experimental source: liver
C;Superfamily: transferrin; transferrin repeat homology
C;Keywords: iron binding

Query Match 28.7%; Score 76.5; DB 2; Length 690;

Search completed: August 11, 2003, 08:18:58
Job time : 14.0485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:16:36 ; Search time 6.9029 Seconds
(without alignments)
322.917 Million cell updates/sec

Title: US-09-508-095-19

Perfect score: 267

Sequence: 1 GRRSVQWCAVSQPEATKCF.....VRGPPVSCIKRDSPIQICQA 48

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	261	97.8	711	1	TRFL_HUMAN
2	164	61.4	695	1	TRFL_HORSE
3	163	61.0	708	1	TRFL_CAPHI
4	162	60.7	708	1	TRFL_CAMDR
5	159	59.6	707	1	TRFL_MOUSE
6	136	50.9	708	1	TRFL_MOUSE
7	134	50.2	708	1	TRFL_BUBBU
8	132.5	49.6	704	1	TRFL_BOVIN
9	120	44.9	696	1	ICA_PIG
10	118.5	44.4	698	1	TRFE_PIG
11	118	44.2	704	1	TRFE_HUMAN
12	112	41.9	704	1	TRFE_PIG
13	110.5	41.4	695	1	TRFE_BOVIN
14	106.5	39.9	697	1	TRFE_RABIT
15	106.5	39.9	698	1	TRFE_MOUSE
16	103	38.6	706	1	TRFE_RAT
17	85.5	32.0	687	1	TRFE_HORSE
18	77.5	29.0	686	1	TRFE_ONCKI
19	76.5	28.7	690	1	TRFE_ANAPL
20	76.5	28.7	691	1	TRFE_SALSA
21	76.5	28.7	738	1	TRFE_SALSA
22	76	28.5	685	1	TRFE_HUMAN
23	76	28.5	705	1	TRFE_PAROL
24	75.5	28.3	738	1	TRFE_CHICK
25	73.5	27.5	690	1	TRFE_MOUSE
26	73	27.3	844	1	TRFE_ORYLA
27	72	27.0	701	1	SAX_RANCA
28	57	21.3	373	1	TRFE_XENIA
29	57	21.3	546	1	TYRA_ERWHE
30	57	21.3	642	1	YTER_CAEEL
31	55	20.6	2505	1	TRFE_GADMO
32	54	20.2	422	1	FAS_RAT
33	53.5	20.0	1371	1	U183_CAEEL
				1	VCAP_HSVSA

34 53 19.9 137 1 CS11_HUMAN
35 53 19.9 373 1 TYRA_ECOLI
36 52.5 19.7 64 1 BD01_CAPHI
37 52.5 19.7 738 1 I12R_MOUSE
38 52.5 19.7 2130 1 BA2B_CHICK
39 52 19.5 333 1 AMP_IMPEA
40 51.5 19.3 681 1 TRF_MANSE
41 51.5 19.3 811 1 SYQ_SCHPO
42 51.5 19.3 1265 1 MYO6_MOUSE
43 51 19.1 274 1 MOVP_TMOB
44 51 19.1 522 1 POLS_RUBV
45 50.5 18.9 525 1 LAG3_HUMAN

ALIGNMENTS

RESULT 1
TRFL_HUMAN
ID TRFL_HUMAN STANDARD; PRT; 711 AA.
AC P02788; O00756; Q16780; Q16786; Q16789; Q96K24; Q96K25;
AC Q9H123;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;
DE Lactoferrin B; Lactoferrin C].
GN LTF OR LF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=90384839; PubMed=2402455;
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";
RL Nucleic Acids Res. 18:5288-5288(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX Cho Y.Y.;
RA Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.
RN [3]
RP SEQUENCE FROM N.A.
RC Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
RX Conneely O.M.;
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX Liang Q., Jimenez-Flores R., Richardson T.;
RA "Molecular cloning and sequence analysis of human lactoferrin.";
RN [5]
RP Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RX Conneely O.M.;
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX Wei X., Han J., Rado T.A.;
RA "Human neutrophil lactoferrin coding and 5' flanking region DNA
RN [7]
RP Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RX Conneely O.M.;
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX Cheng H., Chen X., Huan L.;
RA "cDNA cloning and sequence analysis of human lactoferrin.";
RN [9]
RP Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RX Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RN SEQUENCE OF 3-711 FROM N.A.
RP TISSUE-Mammary gland;
RX MEDLINE=90326549; PubMed=2374734;
RA Powell M.J., Ogden J.E.;
RT "Nucleotide sequence of human lactoferrin cDNA";
RL Nucleic Acids Res. 18:4013-4013(1990).
RN [9]
RN SEQUENCE OF 20-711.
RX MEDLINE=85076667; PubMed=65104020;
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,
RA Legrand D., Spik G., Montreuil J., Jolles P.;
RT "Human lactotransferrin: amino acid sequence and structural
RT comparisons with other transferrins";
RL Eur. J. Biochem. 145:659-666(1984).
RN [10]
RN PRELIMINARY SEQUENCE OF 20-73: 134-171; 257-278 AND 347-530.
RX MEDLINE=82046817; PubMed=6794640;
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,
RA Jolles P.;
RT "The present state of the human lactotransferrin sequence. Study and
RT alignment of the cyanogen bromide fragments and characterization of
RT N- and C-terminal domains";
RL Biochim. Biophys. Acta 670:243-254(1981).
RN [11]
RN SEQUENCE OF 609-711.
RX MEDLINE=82262043; PubMed=7049727;
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,
RA Jolles P.;
RT "An 88 amino acid long C-terminal sequence of human
RT lactotransferrin";
RL FEBS Lett. 142:1107-110(1982).
RN [12]
RN SEQUENCE OF 436-711 FROM N.A.
RX MEDLINE=88001031; PubMed=3477300;
RA Rado T.A., Wei X., Benz E.J. Jr.;
RT "Isolation of lactoferrin cDNA from a human myeloid library and
RT expression of mRNA during normal and leukemic myelopoiesis";
RL Blood 70:989-993(1987).
RN [13]
RN SEQUENCE OF 237-711 FROM N.A.
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,
RA Dragan Y., Gialalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,
RA Sagripanti J.L.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [14]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.
RX MEDLINE=90064528; PubMed=2585506;
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;
RT "Structure of human lactoferrin: crystallographic structure analysis
RT and refinement at 2.8-A resolution.";
J. Mol. Biol. 209:711-734(1989).
RN [15]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RA Haridas M., Anderson B.F., Baker E.N.;
RT "Structure of human diferric lactoferrin refined at 2.2-A
RT resolution.";
RL Acta Crystallogr. D 51:629-646(1995).
RN [16]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.
RX MEDLINE=917156796; PubMed=9003186;
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,
RA Baker E.N.;
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron
RT binding properties and crystal structure of the histidine-
RT 253--methionine mutant.";
RL Biochemistry 36:341-346(1997).
RN [17]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=99190892; PubMed=10089347;
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;
RT "Structure of recombinant human lactoferrin expressed in Aspergillus
RT awamori.";
RL Acta Crystallogr. D 55:403-407(1999).
RN [18]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=99192677; PubMed=10089508;
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement
RT and analysis of ligand-induced conformational change.";
RL Acta Crystallogr. D 54:1319-1335(1998).
RN [19]
RN CHARACTERIZATION OF LACTOFERROXINS.
RX MEDLINE=911166929; PubMed=1369293;
RA Tani F., Ito K., Chiba K., Yoshikawa M.;
RT "Isolation and characterization of opioid antagonist peptides derived
RT from human lactoferrin.";
RL Agric. Biol. Chem. 54:1803-1810(1990).
RN [20]
RN VARIANTS THR-30 AND ARG-48.
RX MEDLINE=99091914; PubMed=9873069;
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,
RA Qumsiyeh W.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,
RA Sagar J., Kumaranickavel G., Munier F., Schorderet D.F.,
RA El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,
RA Hejtmancik J.F., Teng C.T.;
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";
RL Mol. Vision 4:31-32(1998).
RN [21]
RN FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53961; CAA37914.1; -
DR EMBL: U07643; AAB60324.1; -
DR EMBL: M93150; AAA36159.1; -
DR EMBL: M83202; AAA59311.1; -
DR EMBL: M83205; AAA58656.1; -

```

Query Match          97.8%; Score 261; DB 1; Length 711;
Best Local Similarity 100.0%; Pred.No. 3.7e-26;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRSVQWCAVSQPEATKCFQQRNNRKVRGPPVSCIKRDSPIQCIQA 48
   |||||
Db 22 RRRSVQWCAVSQPEATKCFQQRNNRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 2
TRFL_HORSE
ID TRFL_HORSE STANDARD; PRT; 695 AA.
AC O77811;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin) (Fragment).
GS LTF.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "cDNA sequence of mare lactoferrin.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RC TISSUE=Milk;
RX MEDLINE=99296631; PubMed=10366507;
RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;
RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A
RT resolution.";
RL J. Mol. Biol. 289:303-317(1999).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AJ010930; CAA09407.1; -.
PDB; 1BLX; 02-DEC-98.
PDB; 1B7U; 02-FEB-99.
PDB; 1B7Z; 02-FEB-99.
PDB; 1F9B; 10-FEB-01.
PDB; 1I6B; 13-FEB-02.
PDB; 1QJM; 14-JAN-00.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
Signal; 3D-structure.
FT NON_TER 1
FT SIGNAL <1 6
FT CHAIN 7 695 LACTOTRANSFERRIN.
FT REPEAT 7 350 1.
FT REPEAT 351 695 2.

```

```

FT DISULFID 15
FT DISULFID 25
FT DISULFID 121
FT DISULFID 163
FT DISULFID 179
FT DISULFID 189
FT DISULFID 176
FT DISULFID 237
FT DISULFID 354
FT DISULFID 364
FT DISULFID 377
FT DISULFID 411
FT DISULFID 431
FT DISULFID 463
FT DISULFID 487
FT DISULFID 497
FT DISULFID 511
FT DISULFID 508
FT DISULFID 579
FT DISULFID 593
FT DISULFID 631
FT METAL 66
FT METAL 98
FT METAL 198
FT METAL 259
FT METAL 401
FT METAL 439
FT METAL 532
FT METAL 601
FT BINDING 127
FT BINDING 127
FT CARBOHYD 469
FT CARBOHYD 143
FT CARBOHYD 287
FT CARBOHYD 482
FT STRAND 12
FT HELIX 19
FT TURN 35
FT STRAND 40
FT STRAND 44
FT HELIX 48
FT TURN 57
FT STRAND 62
FT HELIX 67
FT TURN 76
FT STRAND 80
FT STRAND 88
FT STRAND 95
FT STRAND 97
FT TURN 106
FT HELIX 112
FT TURN 116
FT STRAND 119
FT TURN 125
FT TURN 126
FT TURN 128
FT HELIX 132
FT HELIX 139
FT TURN 147
FT HELIX 151
FT STRAND 161
FT TURN 165
FT TURN 169
FT HELIX 173
FT TURN 176
FT TURN 183
FT TURN 191
FT TURN 194
FT HELIX 197
FT TURN 207
FT STRAND 212
FT HELIX 219
FT HELIX 227
FT TURN 231
FT STRAND 233
FT TURN 238
FT STRAND 241
FT HELIX 245
FT TURN 248
FT STRAND 254
FT STRAND 260

```

```

IRON 1 (BY SIMILARITY).
IRON 1 (BY SIMILARITY).
IRON 1 (BY SIMILARITY).
IRON 2 (BY SIMILARITY).
IRON 2 (BY SIMILARITY).
IRON 2 (BY SIMILARITY).
IRON 2 (BY SIMILARITY).
ANION (BY SIMILARITY).
ANION (BY SIMILARITY).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

```


DR	Pfam: PF00405; transferrin; 2.
DR	PRINTS: PRQ0422; TRANSFERRIN.
DR	SMART: SMO0094; TR_FER; 2.
DR	PROSITE: PS00205; TRANSFERRIN_1; 1.
DR	PROSITE: PS00206; TRANSFERRIN_2; 2.
DR	PROSITE: PS00207; TRANSFERRIN_3; 2.
KW	Transposite; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW	Signal.
FT	SIGNAL 1 19 BY SIMILARITY.
FT	CHAIN 20 707 LACTOTRANSFERRIN.
FT	REPEAT 20 357 1.
FT	REPEAT 358 707 2.
FT	DISULFID 27 63 BY SIMILARITY.
FT	DISULFID 37 54 BY SIMILARITY.
FT	DISULFID 133 216 BY SIMILARITY.
FT	DISULFID 175 191 BY SIMILARITY.
FT	DISULFID 188 199 BY SIMILARITY.
FT	DISULFID 249 263 BY SIMILARITY.
FT	DISULFID 366 398 BY SIMILARITY.
FT	DISULFID 376 389 BY SIMILARITY.
FT	DISULFID 423 702 BY SIMILARITY.
FT	DISULFID 443 665 BY SIMILARITY.
FT	DISULFID 475 550 BY SIMILARITY.
FT	DISULFID 499 693 BY SIMILARITY.
FT	DISULFID 509 523 BY SIMILARITY.
FT	DISULFID 520 533 BY SIMILARITY.
FT	DISULFID 591 605 BY SIMILARITY.
FT	DISULFID 643 648 BY SIMILARITY.
FT	METAL 78 78 IRON 1 (BY SIMILARITY).
FT	METAL 110 110 IRON 1 (BY SIMILARITY).
FT	METAL 210 210 IRON 1 (BY SIMILARITY).
FT	METAL 271 271 IRON 1 (BY SIMILARITY).
FT	METAL 413 413 IRON 2 (BY SIMILARITY).
FT	METAL 451 451 IRON 2 (BY SIMILARITY).
FT	METAL 544 544 IRON 2 (BY SIMILARITY).
FT	METAL 613 613 IRON 2 (BY SIMILARITY).
FT	BINDING 139 139 ANION (POTENTIAL).
FT	BINDING 481 481 ANION (POTENTIAL).
FT	CARBHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBHYD 494 494 N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CONFLICT 1 2 MR -> IQG (IN REF. 1).
FT	CONFLICT 25 25 R -> Q (IN REF. 2).
FT	CONFLICT 82 82 M -> L (IN REF. 2).
FT	CONFLICT 359 359 S -> T (IN REF. 1).
FT	CONFLICT 382 382 A -> D (IN REF. 1).
FT	CONFLICT 449 449 E -> G (IN REF. 2).
FT	CONFLICT 629 629 L -> V (IN REF. 1).
SQ	SEQUENCE 707 AA; 77865 MW; F26AE0340AC19A8 CRC64; Query Match 59.6%; Score 159; DB 1; Length 707; Best Local Similarity 61.4%; Pred.No.5.e-13; Matches 27; Conservative 6; Mismatches 11; Indels 0;
QY	5 SVOWCAVSOPKATCFQWQNMRKVGGPPVSCIKRDSPICQIA 48 ::: :: :: ::
Db	23 TVRWCAVNSEEEKLRWQMNRKVGGPPPLSKSSSTQCIOIA 66 ::: :: :: ::
RESULT 6	
TREFL_BUBBU	BUBBU
ID TREFL_BUBBU	STANDARD; PRT; 708 AA.
AC 077638;	
DT 15-JUL-1999 (Rel. 38, Created)	
DT 15-JUL-1999 (Rel. 38, Last sequence update)	
DT 15-SEP-2003 (Rel. 42, Last annotation update)	
DE Lactotransferrin precursor (Lactoferrin).	
GN LRF	
OS Bubalus bubalis (Domestic water buffalo).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoid	
OC Bovidae; Bovinae; Bubalus.	
OX NCBI_TaxID=89462;	
RN [1]	

```

RP SEQUENCE FROM N:A.
RA Paramasivam M., Thattaliyath B.D., Kumar A., Srinivasan A.,
RT Singh T.P.;
RL "cDNA sequence of Buffalo lactoferrin.";
RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
(2)
RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RP MEDLINE=20003130; PubMed=10531476;
RA Karthikeyan S., Paramasivam M., Yedav S., Srinivasan A., Singh T.P.;
RT "Structure of Buffalo lactoferrin at 2.5-A resolution using crystals
RN grown at 303 K shows different orientations of the N and C lobes.";
RL Acta Crystallogr. D 55:1805-1813(1999).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ005203; CAA06441.1; -.
DR PDB; 1CE2; 19-MAR-99.
DR PDB; 1BIY; 13-JAN-99.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 708 LACTOTRANSFERRIN.
FT DISULFID 28 64
FT DISULFID 38 55
FT DISULFID 134 217
FT DISULFID 176 192
FT DISULFID 179 202
FT DISULFID 189 200
FT DISULFID 250 264
FT DISULFID 367 399
FT DISULFID 377 390
FT DISULFID 424 703
FT DISULFID 444 666
FT DISULFID 476 551
FT DISULFID 500 694
FT DISULFID 510 524
FT DISULFID 521 534
FT DISULFID 592 606
FT DISULFID 644 649
FT METAL 79 79 IRON 1.
FT METAL 111 111 IRON 1.
FT METAL 211 211 IRON 1.
FT METAL 272 272 IRON 1.
FT METAL 414 414 IRON 2.
FT METAL 452 452 IRON 2.
FT METAL 545 545 IRON 2.
FT METAL 614 614 IRON 2.
FT BINDING 140 140 ANION (POTENTIAL).
FT BINDING 482 482 ANION (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TURN 21 22
FT STRAND 25 29
FT HELIX 32 46
FT TURN 47 48
FT STRAND 53 57
FT HELIX 61 69
FT TURN 70 71
FT STRAND 75 78
FT HELIX 80 87
FT TURN 89 91
FT STRAND 93 102
FT STRAND 107 108
FT STRAND 110 118
FT TURN 119 120
FT HELIX 125 127
FT TURN 129 130
FT STRAND 132 135
FT STRAND 138 139
FT TURN 141 144
FT HELIX 145 150
FT TURN 151 151
FT HELIX 152 155
FT TURN 159 161
FT HELIX 164 169
FT TURN 170 171
FT STRAND 174 176
FT TURN 178 179
FT TURN 182 184
FT HELIX 186 189
FT TURN 190 191
FT HELIX 196 198
FT TURN 199 200
FT TURN 204 205
FT TURN 207 208
FT HELIX 211 219
FT TURN 220 221
FT STRAND 225 229
FT TURN 230 231
FT HELIX 232 236
FT HELIX 240 243
FT TURN 244 245
FT STRAND 246 249
FT TURN 251 252
FT STRAND 255 257
FT HELIX 258 260
FT TURN 261 263
FT STRAND 267 270
FT STRAND 273 277
FT HELIX 283 297
FT TURN 299 300
FT TURN 307 308
FT TURN 312 313
FT TURN 321 322
FT STRAND 325 328
FT TURN 331 332
FT HELIX 335 351
FT STRAND 354 362
FT STRAND 364 369
FT HELIX 371 384
FT TURN 385 385
FT STRAND 388 393
FT TURN 396 404
FT TURN 405 406
FT STRAND 410 413
FT HELIX 415 422
FT TURN 423 425
FT STRAND 426 434
FT TURN 438 441
FT TURN 444 446
FT STRAND 452 459
FT TURN 460 461
FT TURN 463 464
FT HELIX 467 469

```


GN TF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS TF*E2; TF*CHI AND TF*DI.
RX MEDLINE-84194084; PubMed-6585826;
RA Yang F., Lum J.B., McGill J.R., Moore C.M., Naylor S.L.,
RA van Bragt P.H., Baldwin W.D., Bowman B.H.;
RT "Human transferrin: cDNA characterization and chromosomal
RT localization.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:2752-2756(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-88056305; PubMed-3678832;
RA Schaeffer E., Lucero M.A., Jeltsch J.-M., Py M.-C., Levin M.J.,
RA Chambon P., Cohen G.N., Zakin M.M.;
RT "Complete structure of the human transferrin gene. Comparison with
RT analogous chicken gene and human pseudogene.";
RL Gene 56:109-116(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-92231399; PubMed-1809186;
RA Hersberger C.L., Larson J.L., Arnold B., Rosteck P.R. Jr.,
RA Williams P., Dehoff B., Dunn P., O'Neal K.L., Riemen M.W.,
RA Tice P.A.;
RT "A cloned gene for human transferrin.";
RL Ann. N.Y. Acad. Sci. 646:140-154(1991).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT ATRANSFERRINEMIA PRO-477.
RX MEDLINE-20563920; PubMed-11110675;
RA Beutler E., Gelbart T., Lee P.L., Trevino R., Fernandez M.A.,
RA Fairbanks V.F.;
RT "Molecular characterization of a case of atransferrinemia.";
RL Blood 96:4071-4074(2000).
RN [5]
RP SEQUENCE OF 99-698 FROM N.A.
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Bi J., Zhang Y., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 33 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 422-698 FROM N.A.
RX MEDLINE-84153910; PubMed-6322780;
RA Uzan G., Frain M., Park I., Besmond C., Maessen G., Trepas J.S.,
RA Zakin M.M., Kahn A.;
RT "Molecular cloning and sequence analysis of cDNA for human
RT transferrin.";
RL Biochem. Biophys. Res. Commun. 119:273-281(1984).
RN [7]
RP SEQUENCE OF 20-698.
RX MEDLINE-83160878; PubMed-6833213;
RA McGillivray R.T.A., Mendez E., Shewale J.G., Sinha S.K.,
RA Lineback-Zins J., Brew K.;
RT "The primary structure of human serum transferrin. The structures of
RT seven cyanogen bromide fragments and the assembly of the complete
RT structure.";
RL J. Biol. Chem. 258:3543-3553(1983).
RN [8]
RP SEQUENCE OF 73-698 FROM N.A.
RX MEDLINE-85216459; PubMed-3858812;
RA Park I., Schaeffer E., Sidoli A., Baralle F.E., Cohen G.N.,
RA Zakin M.M.;
RT "Organization of the human transferrin gene: direct evidence that it
RT originated by gene duplication.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3149-3153(1985).
RN [9]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE-87066744; PubMed-3786138;

RA Lucero M.A., Schaeffer E., Cohen G.N., Zakin M.M.;
RT "The 5' region of the human transferrin gene: structure and potential
RT regulatory sites.";
RL Nucleic Acids Res. 14:8692-8692(1986).
RN [10]
RP SEQUENCE OF 1-72 AND 291-300 FROM N.A.
RX MEDLINE-87192006; PubMed-3106157;
RA Adrian G.S., Korinek B.W., Bowman B.H., Yang F.;
RT "The human transferrin gene: 5' region contains conserved sequences
RT which match the control elements regulated by heavy metals,
RT glucocorticoids and acute phase reaction.";
RL Gene 49:167-175(1986).
RN [11]
RP SEQUENCE OF 45-72 FROM N.A.
RX MEDLINE-20392111; PubMed-10931525;
RA de Arriba Zepa G.A., Saleh M.-C., Fernandez P.M., Guillou F.,
RA Espinosa de los Monteros A., de Vellis J., Zakin M.M., Baron B.;
RT "Alternative splicing prevents transferrin secretion during
RT differentiation of a human oligodendrocyte cell line.";
RL J. Neurosci. Res. 61:388-395(2000).
RN [12]
RP SEQUENCE OF 564-624 FROM N.A., AND VARIANT TF*C2.
RC TISSUE=Brain;
RX MEDLINE-97418135; PubMed-9272172;
RA Namekata K., Oyama F., Imagawa M., Ihara Y.;
RT "Human transferrin (Tf): a single mutation at codon 570 determines Tf
RT C1 or Tf C2 variant.";
RL Hum. Genet. 100:457-458(1997).
RN [13]
RP SEQUENCE OF 564-624 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE-96007936; PubMed-7498159;
RA Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A.,
RA Ershova E.S., Egorov T.A., Musalyanov A.K.;
RT "The major protein expression profile and two-dimensional protein
RT database of human heart.";
RL Electrophoresis 16:1160-1169(1995).
RN [16]
RP DISULFIDE BONDS.
RX MEDLINE-82221166; PubMed-6953407;
RA McGillivray R.T.A., Mendez E., Sinha S.K., Sutton M.R.,
RA Lineback-Zins J., Brew K.;
RT "The complete amino acid sequence of human serum transferrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2504-2508(1982).
RN [17]
RP MUTAGENESIS.
RX MEDLINE-92031536; PubMed-1932003;
RA Woodworth R.C., Mason A.B., Funk W.D., McGillivray R.T.A.;
RT "Expression and initial characterization of five site-directed
RT mutants of the N-terminal half-molecule of human transferrin.";
RL Biochemistry 30:10824-10829(1991).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 22-350.
RX MEDLINE-98272665; PubMed-9609685;
RA Macgillivray R.T.A., Moore S.A., Chen J., Anderson B.F., Baker H.,
RA Luo Y., Bewley M.C., Smith C.A., Murphy M.E.P., Wang Y., Mason A.B.,
RA Woodworth R.C., Brayer G.D., Baker E.N.;
RT "Two high-resolution crystal structures of the recombinant N-lobe of
RT human transferrin reveal a structural change implicated in iron
RT release.";
RL Biochemistry 37:7919-7928(1998).
RN [19]


```
FT METAL 610 610 IRON 2 (BY SIMILARITY).
FT BINDING 136 136 ANION (BY SIMILARITY).
FT BINDING 478 478 ANION (BY SIMILARITY).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 12 12 G -> W (IN REF. 1).
FT CONFLICT 46 48 RRT -> TTR (IN REF. 3).
FT CONFLICT 51 51 M -> I (IN REF. 1).
FT CONFLICT 85 85 D -> G (IN REF. 1).
FT CONFLICT 121 121 MISSING (IN REF. 2).
FT CONFLICT 132 132 T -> I (IN REF. 1).
FT CONFLICT 284 284 E -> S (IN REF. 1).
FT CONFLICT 573 573 E -> Q (IN REF. 1).
FT CONFLICT 590 590 D -> N (IN REF. 1).
FT CONFLICT 625 625 V -> M (IN REF. 1).
FT CONFLICT 662 662 V -> C (IN REF. 1).
FT CONFLICT 686 704 NLKQSVSPLEACAFMMR -> T (IN REF. 1).
SQ SEQUENCE 704 AA; 77625 MW; 93261EFD608AD358 CRC64;

Query Match 44.2%; Score 118; DB 1; Length 704;
Best Local Similarity 43.5%; Pred. No. 1.1e-07;
Matches 20; Conservative 10; Mismatches 14; Indels 2; Gaps 1;

QY 3 RRSQWCAVSQPEATKCFQWQNRKVRGPPVSCIKRDSPIQIOA 48
DB 22 KKGVRWCVISTAEYSKQWQSKIRRTN--PMFCIRASPTDCIRA 65

RESULT 12
TRFE_BOVIN
ID TRFE_BOVIN STANDARD; PRT; 704 AA.
AC Q29443;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sero transferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal
DE binding globulin).
GN TF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=96132861; PubMed=8557646;
RT "Production and characterization of chimeric transferrins for the
RT determination of the binding domains for bacterial transferrin
RT receptors";
RL J. Biol. Chem. 271:1166-1173(1996).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE. IT IS RESPONSIBLE FOR THE
CC TRANSPORT OF IRON FROM SITES OF ABSORPTION AND HEME DEGRADATION
CC TO THOSE OF STORAGE AND UTILIZATION. SERUM TRANSFERRIN MAY ALSO
CC HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed in liver; secreted in plasma.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02564; AAA96735.1; ..
```

```
DR HSSP; P19134; 1TFD.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW signal.
FT CHAIN 1 19 BY SIMILARITY
FT CHAIN 20 704 SEROTRANSFERRIN.
FT REPEAT 20 359 1.
FT REPEAT 360 704 2.
FT DISULFID 28 66 BY SIMILARITY.
FT DISULFID 38 57 BY SIMILARITY.
FT DISULFID 136 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 179 200 BY SIMILARITY.
FT DISULFID 189 202 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 362 622 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 699 BY SIMILARITY.
FT DISULFID 441 663 BY SIMILARITY.
FT DISULFID 473 549 BY SIMILARITY.
FT DISULFID 497 690 BY SIMILARITY.
FT DISULFID 518 532 BY SIMILARITY.
FT DISULFID 589 603 BY SIMILARITY.
FT DISULFID 641 646 BY SIMILARITY.
FT METAL 81 81 IRON 1 (BY SIMILARITY).
FT METAL 113 113 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 449 449 IRON 2 (BY SIMILARITY).
FT METAL 543 543 IRON 2 (BY SIMILARITY).
FT METAL 611 611 IRON 2 (BY SIMILARITY).
FT BINDING 142 142 ANION (POTENTIAL).
FT BINDING 479 479 ANION (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 704 AA; 77753 MW; D87BB2AFE46C708D CRC64;
```

Query Match 41.9%; Score 112; DB 1; Length 704;
Best Local Similarity 42.6%; Pred. No. 6.5e-07;
Matches 20; Conservative 12; Mismatches 13; Indels 2; Gaps 1;

QY 4 RSVQWCAVSQPEATKCFQWQNRKVR--RGPPVSCIKRDSPIQIOA 48
DB 23 RTVRWCTISTHANKCAFSENVLRILESGPFVSCVKTKTSHMDCIKA 69

RESULT 13
TRFE_RABIT
ID TRFE_RABIT STANDARD; PRT; 695 AA.
AC P19134; O46514;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sero transferrin precursor (transferrin) (Siderophilin) (Beta-1-metal
DE binding globulin).
GN TF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91274362; PubMed=2054387;
RA Banfield D.K., Chow B.K.-C., Funk W.D., Robertson K.A., Umelas T.M.,
RA Woodworth R.C., Macgillivray R.T.A.;

RT "The nucleotide sequence of rabbit liver transferrin cDNA.";
RL Biochim. Biophys. Acta 1089:262-265(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-New Zealand white;
RA Gharab B.A., Thepot D., Puissant C., Cajero-Juarez M.,
RA Houdebine L.M.;
RT "Cloning and structural organisation of the rabbit transferrin encoding
gene.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 20-51.
RX MEDLINE=88209278; PubMed=3365331;
RA Godovac-Zimmermann J.;
RT "Isolation, characterization and N-terminal amino-acid sequence of
rabbit transferrin.";
RL Biol. Chem. Hoppe-Seyler 369:93-96(1988).
RN [4]
RP SEQUENCE OF 483-545.
RX MEDLINE=89005676; PubMed=3169252;
RA Evans R.W., Aitken A., Patel K.J.;
RT "Evidence for a single glycan moiety in rabbit serum transferrin and
location of the glycan within the polypeptide chain.";
RL FEBS Lett. 238:39-42(1988).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RC TISSUE=Plasma;
RX MEDLINE=89026775; PubMed=3179277;
RA Bailey S., Evans R.W., Garratt R.C., Gorinsky B., Hasnain S.,
RA Horsburgh C., Jhoti H., Lindley P.F., Mydin A., Sarra R., Watson J.L.;
RT "Molecular structure of serum transferrin at 3.3-A resolution.";
RL Biochemistry 27:5804-5812(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RA Sarra R., Garratt R.C., Gorinsky B., Jhoti H., Lindley P.F.;
RT "High-resolution X-ray studies on rabbit serum transferrin:
preliminary structure analysis of the N-terminal half-molecule at
2.3-A resolution.";
RL Acta Crystallogr. B 46:763-771(1990).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
OF AN ANION. USUALLY BICARBONATE. IT IS RESPONSIBLE FOR THE
TRANSPORT OF IRON FROM SITES OF ABSORPTION AND HEME DEGRADATION
TO THOSE OF STORAGE AND UTILIZATION. SERUM TRANSFERRIN MAY ALSO
HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed in liver; secreted in plasma.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58533; CAA41424.1; -;
DR EMBL; AF031625; AAB94136.1; JOINED.
DR EMBL; AF031611; AAB94136.1; JOINED.
DR EMBL; AF031612; AAB94136.1; JOINED.
DR EMBL; AF031613; AAB94136.1; JOINED.
DR EMBL; AF031614; AAB94136.1; JOINED.
DR EMBL; AF031615; AAB94136.1; JOINED.
DR EMBL; AF031616; AAB94136.1; JOINED.
DR EMBL; AF031617; AAB94136.1; JOINED.
DR EMBL; AF031618; AAB94136.1; JOINED.
DR EMBL; AF031619; AAB94136.1; JOINED.
DR EMBL; AF031620; AAB94136.1; JOINED.
DR EMBL; AF031621; AAB94136.1; JOINED.
DR EMBL; AF031622; AAB94136.1; JOINED.
DR EMBL; AF031623; AAB94136.1; JOINED.
DR EMBL; AF031624; AAB94136.1; JOINED.
DR PDB; 1TFD; 15-APR-93.
DR PDB; 1JNF; 09-JAN-02.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 695 SEROTRANSFERRIN.
FT REPEAT 20 355 1.
FT REPEAT 356 695 2.
FT DISULFID 28 67 BY SIMILARITY.
FT DISULFID 38 58 BY SIMILARITY.
FT DISULFID 137 213 BY SIMILARITY.
FT DISULFID 156 350 BY SIMILARITY.
FT DISULFID 177 193 BY SIMILARITY.
FT DISULFID 180 196 BY SIMILARITY.
FT DISULFID 190 198 BY SIMILARITY.
FT DISULFID 246 260 BY SIMILARITY.
FT DISULFID 358 612 BY SIMILARITY.
FT DISULFID 364 396 BY SIMILARITY.
FT DISULFID 374 387 BY SIMILARITY.
FT DISULFID 421 690 BY SIMILARITY.
FT DISULFID 436 653 BY SIMILARITY.
FT DISULFID 468 539 BY SIMILARITY.
FT DISULFID 492 681 BY SIMILARITY.
FT DISULFID 502 516 BY SIMILARITY.
FT DISULFID 513 522 BY SIMILARITY.
FT DISULFID 579 593 BY SIMILARITY.
FT DISULFID 631 636 BY SIMILARITY.
FT CARBOHYD 509 509 N-LINKED (GLCNAC...).
FT METAL 82 82 IRON 1 (BY SIMILARITY).
FT METAL 114 114 IRON 1 (BY SIMILARITY).
FT METAL 207 207 IRON 1 (BY SIMILARITY).
FT METAL 268 268 IRON 1 (BY SIMILARITY).
FT METAL 411 411 IRON 2 (BY SIMILARITY).
FT METAL 444 444 IRON 2 (BY SIMILARITY).
FT METAL 533 533 IRON 2 (BY SIMILARITY).
FT METAL 601 601 IRON 2 (BY SIMILARITY).
FT BINDING 143 143 ANION (POTENTIAL).
FT BINDING 474 474 ANION (POTENTIAL).
FT VARIANT 517 517 V -> I.
FT CONFLICT 7 7 MISSING (IN REF. 1).
FT CONFLICT 47 47 K -> S (IN REF. 3).
FT CONFLICT 50 50 P -> Y (IN REF. 3).
FT STRAND 27 29
FT HELIX 32 45
FT TURN 46 48
FT STRAND 59 60
FT HELIX 64 72
FT TURN 73 74
FT STRAND 78 81
FT HELIX 83 90
FT TURN 91 94
FT STRAND 96 102
FT STRAND 114 121
FT TURN 122 123
FT TURN 128 129
FT TURN 132 133
FT TURN 136 137
FT TURN 141 142
FT TURN 144 147
FT HELIX 148 154
FT TURN 155 157
FT HELIX 165 170
FT TURN 171 172
FT STRAND 176 177

```

FT TURN 179 180
FT STRAND 190 190
FT TURN 194 195
FT HELIX 207 214
FT TURN 215 217
FT STRAND 221 225
FT TURN 226 227
FT HELIX 228 232
FT TURN 236 239
FT TURN 240 241
FT STRAND 242 246
FT TURN 247 249
FT STRAND 250 253
FT TURN 254 255
FT TURN 257 259
FT STRAND 263 265
FT STRAND 269 273
FT HELIX 279 288
FT TURN 289 291
FT TURN 303 304
FT STRAND 306 306
FT TURN 307 308
FT STRAND 309 309
FT TURN 316 317
FT STRAND 321 322
SQ SEQUENCE 695 AA; 76670 MW; DB12F34D87AE9D55 CRC64;

Query Match 41.4%; Score 110.5; DB 1; Length 695;
Best Local Similarity 41.7%; Pred. NO. le-06;
Matches 20; Conservative 14; Mismatches 11; Indels 3; Gaps 1;

OY 4 RSVQWCAVSOPEAKCFQWQNRKV---RGPVSCIKRDSPIQICQA 48
DB 23 KTVRCVANDHEASCANFRDSMKVLPEDGPRICVKASYLDCIKA 70

RESULT 14
TREE_MOUSE
ID TREE_MOUSE STANDARD: PRT: 697 AA.
AC Q92111; O35421; Q61803; Q62358; Q62359; Q63915; Q64515; Q8VI15;
AC Q922C0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sero transferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal
DE binding globulin).
GN TF OR TRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c; TISSUE=Liver;
RA Lai D.-Z.;
RT "Construction of a robust CHO cell-line for biopharmaceutical
RT production.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherich A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 1-11 FROM N.A.
RC STRAIN-BALB/c;
RX MEDLINE=96284323; PubMed=9621303;
RA Chaudhary J., Skinner M.K.;
RT "Comparative sequence analysis of the mouse and human transferrin
RT promoters: hormonal regulation of the transferrin promoter in Sertoli
RL cells.";
RL Mol. Reprod. Dev. 50:273-283(1998).
RN [4]
RP SEQUENCE OF 277-337; 462-496 AND 526-575 FROM N.A.
RX MEDLINE=88086992; PubMed=3693348;
RA Chen L.-H., Bissell M.J.;
RT "Transferrin mRNA level in the mouse mammary gland is regulated by
RT pregnancy and extracellular matrix";
RL J. Biol. Chem. 262:17247-17250(1987).
RN [5]
RP SEQUENCE OF 268-307 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94068311; PubMed=8248032;
RA Kasik J.W., Rice E.J.;
RT "Transferrin gene expression in maternal liver, fetal liver and
RT placenta during pregnancy in the mouse.";
RL Placenta 14:365-371(1993).
RN [6]
RP SEQUENCE OF 282-412 FROM N.A.
RX MEDLINE=87280033; PubMed=3611056;
RA Pentecost B.T., Teng C.T.;
RT "Lactotransferrin is the major estrogen inducible protein of mouse
RT uterine secretions.";
RL J. Biol. Chem. 262:10134-10139(1987).
RN [7]
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE. IT IS RESPONSIBLE FOR THE
CC TRANSPORT OF IRON FROM SITES OF ABSORPTION AND HEME DEGRADATION
CC TO THOSE OF STORAGE AND UTILIZATION. SERUM TRANSFERRIN MAY ALSO
CC HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in liver; secreted in plasma.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF440692; AAL34533.1; -
CC EMBL; BC012313; AAH12313.1; -
CC EMBL; BC022986; AAH22986.1; -
CC EMBL; BC008559; AAH08559.1; -
CC EMBL; AF027336; AAB84034.1; -
CC EMBL; M23014; AAB40489.1; -
CC EMBL; M23015; AAA40490.1; -
CC EMBL; M23016; AAA40491.1; -
CC EMBL; S67217; AAB28966.2; -
CC EMBL; J03299; AAA39438.1; -
CC EMBL; A28446; A28446.
CC MGD; MGI:98821; Trf.
CC GO: GO:0030139; C:endocytic vesicle; IDA.

```

```

DR DR InterPro: IPR001156; Transferrin.
DR DR Pfam: PF00405; Transferrin; 2.
DR DR PRINTS: PR00422; TRANSFERRIN.
DR DR PROSITE: PS00205; TRANSFERRIN_1; 1.
DR DR PROSITE: PS00206; TRANSFERRIN_2; 2.
DR DR PROSITE: PS00207; TRANSFERRIN_3; 2.
DR DR Signal:
KW KW Signal; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW KW Signal.
FT FT SIGNAL 1 19 BY SIMILARITY.
FT FT CHAIN 20 697 SEROTRANSFERRIN.
FT FT REPEAT 20 355 1.
FT FT REPEAT 356 697 2.
FT FT DISULFID 28 67 BY SIMILARITY.
FT FT DISULFID 38 58 BY SIMILARITY.
FT FT DISULFID 137 213 BY SIMILARITY.
FT FT DISULFID 156 350 BY SIMILARITY.
FT FT DISULFID 177 193 BY SIMILARITY.
FT FT DISULFID 180 196 BY SIMILARITY.
FT FT DISULFID 190 198 BY SIMILARITY.
FT FT DISULFID 246 260 BY SIMILARITY.
FT FT DISULFID 363 395 BY SIMILARITY.
FT FT DISULFID 373 386 BY SIMILARITY.
FT FT DISULFID 420 692 BY SIMILARITY.
FT FT DISULFID 435 655 BY SIMILARITY.
FT FT DISULFID 472 543 BY SIMILARITY.
FT FT DISULFID 496 683 BY SIMILARITY.
FT FT DISULFID 506 520 BY SIMILARITY.
FT FT DISULFID 517 526 BY SIMILARITY.
FT FT DISULFID 583 597 BY SIMILARITY.
FT FT DISULFID 633 638 BY SIMILARITY.
FT FT METAL 82 82 IRON 1 (BY SIMILARITY).
FT FT METAL 114 114 IRON 1 (BY SIMILARITY).
FT FT METAL 207 207 IRON 1 (BY SIMILARITY).
FT FT METAL 268 268 IRON 2 (BY SIMILARITY).
FT FT METAL 410 410 IRON 2 (BY SIMILARITY).
FT FT METAL 448 448 IRON 2 (BY SIMILARITY).
FT FT METAL 537 537 IRON 2 (BY SIMILARITY).
FT FT METAL 605 605 ANTON (POTENTIAL).
FT FT BINDING 143 143 ANTON (POTENTIAL).
FT FT BINDING 477 477 ANTON (POTENTIAL).
FT FT CARBOHYD 513 513 LT -> FA (IN REF. 1).
FT FT CONFLICT 3 4 ISAS -> HASG (IN REF. 2; AA08559).
FT FT CONFLICT 71 74 W -> L (IN REF. 5).
FT FT CONFLICT 283 283 P -> L (IN REF. 5).
FT FT CONFLICT 307 307 CP -> SA (IN REF. 6).
FT FT CONFLICT 350 351 G -> C (IN REF. 4).
FT FT CONFLICT 487 487 A -> D (IN REF. 4).
FT FT CONFLICT 527 527 K -> N (IN REF. 4).
FT FT CONFLICT 575 575 H -> S (IN REF. 1).
FT FT CONFLICT 697 697
SQ SEQUENCE 697 AA; 76723 MW; 0996A0C3B64CB1B9 CRC64;

Query Match 39.9%; Score 106.5; DB 1; Length 697;
Best Local Similarity 41.7%; Pred. No. 3.3e-06;
Matches 20; Conservative 13; Mismatches 12; Indels 3; Gaps 1;

QY 4 RSYQWCAVSPQATKCFQWQNNKRV---RGPPVSCIKRDSPTQICQA 48
Db 23 KTVKCAVSEHENTKICISFDHMKTVLPDGPRLACVKKTSYDPCIKA 70

RESULT 15
TREE_RAT
ID TREE_RAT STANDARD; PRT; 698 AA.
AC P12346; Q64528; Q64630;
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serotransferrin precursor (transferrin) (Siderophilin) (Beta-1-metal
DE binding globulin).
GN TF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Wistar; TISSUE=Liver;
RX MEDLINE=96208933; PubMed=8829802;
RA Hosino A., Hisayasu S., Shimada T.;
RT "Complete sequence analysis of rat transferrin and expression of
RT transferrin but not lactoferrin in the digestive glands.";
RL Comp. Biochem. Physiol. 113B:491-497(1996).
RN RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Wistar; TISSUE=Mammary gland;
RX MEDLINE=95234054; PubMed=7717992;
RA Escriva H., Pierce A., Coddeville B., Gonzalez F., Benaissa M.,
RA Leger D., Wierzeski J.M., Spik G., Pamblanco M.;
RT "Rat mammary-gland transferrin: nucleotide sequence, phylogenetic
RT analysis and glycan structure.";
RL Biochem. J. 307:47-55(1995).
RN RN [3]
RP SEQUENCE OF 20-47.
RX MEDLINE=80049855; PubMed=500689;
RA Schreiber G., Dryburgh H., Millership A., Matsuda Y., Inglis A.,
RA Phillips J., Edwards K., Maggs J.;
RT "The synthesis and secretion of rat transferrin.";
RL J. Biol. Chem. 254:12013-12019(1979).
RN RN [4]
RP SEQUENCE OF 521-698 FROM N.A.
RX MEDLINE=87053639; PubMed=3023031;
RA Huggenvik J.I., Idzerda R.L., Haywood L., Lee D.C., McKnight G.S.,
RA Griswold M.D.;
RT "Transferrin messenger ribonucleic acid: molecular cloning and
RT hormonal regulation in rat sertoli cells.";
RL Endocrinology 120:332-340(1987).
CC CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE. IT IS RESPONSIBLE FOR THE
CC TRANSPORT OF IRON FROM SITES OF ABSORPTION AND HEME DEGRADATION
CC TO THOSE OF STORAGE AND UTILIZATION. SERUM TRANSFERRIN MAY ALSO
CC HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in liver; secreted in plasma.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR DR EMBL: D38380; BAA07458.1; -
DR DR EMBL: X77158; CAA54403.1; -
DR DR EMBL: M27966; AAA42267.1; -
DR DR PIR: S49163; S49163.
DR DR HSP: P19134; 1TFD.
DR DR GlycoSuiteDB: P12346; -
DR DR InterPro: IPR001156; Transferrin.
DR DR Pfam: PF00405; transferrin; 2.
DR DR PRINTS: PR00422; TRANSFERRIN.
DR DR PROSITE: PS00205; TRANSFERRIN_1; 1.
DR DR PROSITE: PS00206; TRANSFERRIN_2; 2.
DR DR PROSITE: PS00207; TRANSFERRIN_3; 2.
KW KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW KW SIGNAL.
FT FT SIGNAL 1 19
FT FT CHAIN 20 698 SEROTRANSFERRIN.
FT FT REPEAT 20 355 1.
FT FT REPEAT 356 698 2.
FT FT DISULFID 28 67 BY SIMILARITY.

```

```
FT DISULFID 38 58 BY SIMILARITY.
FT DISULFID 137 213 BY SIMILARITY.
FT DISULFID 156 350 BY SIMILARITY.
FT DISULFID 177 193 BY SIMILARITY.
FT DISULFID 180 196 BY SIMILARITY.
FT DISULFID 190 198 BY SIMILARITY.
FT DISULFID 246 260 BY SIMILARITY.
FT DISULFID 263 395 BY SIMILARITY.
FT DISULFID 373 386 BY SIMILARITY.
FT DISULFID 420 693 BY SIMILARITY.
FT DISULFID 435 656 BY SIMILARITY.
FT DISULFID 471 542 BY SIMILARITY.
FT DISULFID 495 684 BY SIMILARITY.
FT DISULFID 505 519 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 634 639 BY SIMILARITY.
FT METAL 82 82 IRON 1 (BY SIMILARITY).
FT METAL 114 114 IRON 1 (BY SIMILARITY).
FT METAL 207 207 IRON 1 (BY SIMILARITY).
FT METAL 268 268 IRON 1 (BY SIMILARITY).
FT METAL 410 410 IRON 2 (BY SIMILARITY).
FT METAL 447 447 IRON 2 (BY SIMILARITY).
FT METAL 536 536 IRON 2 (BY SIMILARITY).
FT METAL 604 604 IRON 2 (BY SIMILARITY).
FT BINDING 143 143 ANION (POTENTIAL).
FT BINDING 477 477 ANION (POTENTIAL).
FT CARBOHYD 512 512 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 57 57 A -> P (IN REF. 2).
FT CONFLICT 110 110 P -> R (IN REF. 2).
FT CONFLICT 318 354 AFGCYGVPDMYRLYLGLGHSYVTAIRNOREGVCPEAS ->
REFLLRAPKGLQAVPRQLCHCHSKSAGSCPDA (IN
REF. 2).
FT CONFLICT 380 380 S -> G (IN REF. 2).
FT CONFLICT 691 691 E -> D (IN REF. 4).
FT CONFLICT 696 697 HK -> TA (IN REF. 4).
SQ SEQUENCE 698 AA; 76363 MW; DDF2C1918E2A1B0E CRC64;

Query Match 39.9%; Score 106.5; DB 1; Length 698;
Best Local Similarity 41.7%; Pred. No. 3.3e-06;
Matches 20; Conservative 13; Mismatches 12; Indels 3; Gaps 1;

OY 4 RSVQWCAVSQPEATKCFOWQNRMRKV---RGPPVSCIKRDSPIOCIOA 48
:::|||||: |||::: |: | ||::|: | ||:|
DB 23 KTVKCAVSEHENTKCSFRDHMKTVLPADGPRLCVAKTSYQDCIKA 70
```

Search completed: August 11, 2003, 08:17:01
Job time : 7.99029 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:16:37 ; Search time 27.9612 Seconds
(without alignments)
442.990 Million cell updates/sec

Title: US-09-508-095-19
Perfect score: 267
Sequence: 1 GRRSVQWCAVSQPEATKCF.....VRGPPVSCIKRDSPIQCIQA 48

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	261	97.8	234	Q8IX02	Q8IX02 homo sapien
2	254	95.1	711	Q8IZH6	Q8IZH6 homo sapien
3	254	95.1	711	Q8IU92	Q8IU92 homo sapien
4	253	94.8	711	Q8TCD2	Q8TCD2 homo sapien
5	163	61.0	707	Q8CBA0	Q8CBA0 mus musculus
6	144.5	54.1	38	Q8UCI5	Q8UCI5 mus musculus
7	120	44.9	704	Q8WMN8	Q8WMN8 sus scrofa
8	115	43.1	33	Q8TR80	Q8TR80 ovis aries
9	112.5	42.1	700	Q8DBD0	Q8DBD0 mus musculus
10	106.5	39.9	700	Q8VC96	Q8VC96 mus musculus
11	103	38.6	54	Q8TQV7	Q8TQV7 equus caball
12	101	37.8	711	Q8XT72	Q8XT72 trichosurus
13	92.5	34.6	87	Q63602	Q63602 rattus norv
14	86.5	32.4	691	Q9PT13	Q9PT13 oncorhynch
15	85.5	32.0	517	Q8UUK7	Q8UUK7 oncorhynch
16	85.5	32.0	672	Q9PU08	Q9PU08 oncorhynch

17	83.5	31.3	382	13	Q9DFK6	Q9DFK6 gillichthys
18	82.5	30.9	677	13	Q9I8R0	Q9I8R0 oncorhynch
19	82.5	30.9	677	13	Q9PT58	Q9PT58 oncorhynch
20	82.5	30.9	677	13	Q9I866	Q9I866 oncorhynch
21	81.5	30.5	676	13	Q9I8R2	Q9I8R2 oncorhynch
22	81.5	30.5	677	13	Q9PT53	Q9PT53 oncorhynch
23	81.5	30.5	677	13	Q9PT55	Q9PT55 oncorhynch
24	81.5	30.5	677	13	Q9I8R3	Q9I8R3 oncorhynch
25	81.5	30.5	677	13	Q9I887	Q9I887 oncorhynch
26	80.5	30.1	367	13	Q9IBF7	Q9IBF7 oncorhynch
27	79.5	29.8	691	13	Q9PU70	Q9PU70 oncorhynch
28	78.5	29.4	72	13	Q8AUN8	Q8AUN8 salmo trutt
29	78.5	29.4	691	13	Q9PU66	Q9PU66 salmo trutt
30	77.5	29.0	691	13	Q9PRH5	Q9PRH5 oncorhynch
31	77.5	29.0	691	13	Q9PU67	Q9PU67 salvelinus
32	76.5	28.7	71	13	Q8AUP0	Q8AUP0 salmo salar
33	76.5	28.7	672	13	Q9PT57	Q9PT57 oncorhynch
34	76.5	28.7	677	13	Q9I858	Q9I858 oncorhynch
35	76.5	28.7	677	13	Q9I8Q9	Q9I8Q9 oncorhynch
36	76.5	28.7	691	13	Q9PU68	Q9PU68 salvelinus
37	76	28.5	686	13	Q9IA82	Q9IA82 paralichthy
38	75.5	28.3	71	13	Q8AQQ2	Q8AQQ2 salmo trutt
39	75.5	28.3	71	13	Q8AUQ1	Q8AUQ1 salmo trutt
40	75.5	28.3	71	13	Q8AUQ0	Q8AUQ0 salmo trutt
41	75.5	28.3	72	13	Q8AYF8	Q8AYF8 salmo trutt
42	75.5	28.3	72	13	Q8AUN7	Q8AUN7 salmo trutt
43	75.5	28.3	672	13	Q9PT56	Q9PT56 oncorhynch
44	75.5	28.3	677	13	Q9I860	Q9I860 oncorhynch
45	75.5	28.3	677	13	Q9PT54	Q9PT54 oncorhynch

ALIGNMENTS

RESULT 1

Q8IX02 ID Q8IX02 PRELIMINARY; PRT; 234 AA.
AC Q8IX02, 2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Lactoferrin (Fragment)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Z., Li N.
RT "Human lactoferrin gene, 5' regulation region and exons 1-6."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF508798; AA014686.1; -
FT NON_TER 234 234
SQ SEQUENCE 234 AA; 25671 MW; 2A08DAFD6B8780C7 CRC64;

Query Match 97.8%; Score 261; DB 4; Length 234;
Best Local Similarity 100.0%; Pred. NO. le-28;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RRRSVQWCAVSQPEATKCFQWQNRKVRGPPVSCIKRDSPIQCIQA 48
|||||
Db 22 RRRSVQWCAVSQPEATKCFQWQNRKVRGPPVSCIKRDSPIQCIQA 68

RESULT 2

Q8IZH6 ID Q8IZH6 PRELIMINARY; PRT; 711 AA.
AC Q8IZH6, 2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Lactoferrin (Fragment)
OS Homo sapiens (Human).

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Kaplan J.B., Fine D.H.;
RT	"Characterization of an amino acid polymorphism in the antibacterial
RT	domain of human lactoferrin.";
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY137470; AAN11304.1; -
SQ	SEQUENCE 711 AA; 78396 MW; 547AB9423C27CE67 CRC64;
QY	2 RRRSVQWCAVSOPATKCFQWRNMKRVGPPVSCIKRDSPIQCIA 48
Dy	22 RRRSVQWCTVSOPATKCFQWRNMKRVGPPVSCIKRDSPIQCIA 68
RESULT 3	
O8IU92	
ID	Q8IU92 PRELIMINARY; PRT; 711 AA.
AC	O8IU92;
DT	01-MAR-2003 (TReMBLrel. 23; Created)
DT	01-MAR-2003 (TReMBLrel. 23; Last sequence update)
DT	01-MAR-2003 (TReMBLrel. 23; Last annotation update)
DE	Lactotransferrin precursor (Lactoferrin).
GN	LTF.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Mammary gland;
RA	Shi Y.-Q.; Zhang Y.; Zheng Y.-M.;
RT	"Homo sapiens Lactotransferrin Gene: cDNA Cloning and Sequence
RT	Analysis.";
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Seminial vesicle;
RA	Baskar Singh S., Saravanan K., Paramasivam M., Srinivasan A.,
RA	Singh T.P.;
RT	"Homo sapiens lactoferrin (HLF) mRNA.";
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY15046; AAN63998.1; -
DR	EMBL; AY178998; AAN75578.2; -
KW	Signal.
FT	SIGNAL
FT	CHAIN
SQ	SEQUENCE 711 AA; 78382 MW; 547BFC42C9267E67 CRC64;
Query Match	95.1%; Score 254; DB 4; Length 711;
Best Local Similarity	95.7%; Pred. No. 3.le-27;
Matches	45; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY	2 RRRSVQWCAVSOPATKCFQWRNMKRVGPPVSCIKRDSPIQCIA 48
Dy	22 RRRSVQWCTVSOPATKCFQWRNMKRVGPPVSCIKRDSPIQCIA 68
RESULT 4	
O8TCD2	
ID	Q8TCD2 PRELIMINARY; PRT; 711 AA.
AC	O8TCD2;
DT	01-JUN-2002 (TReMBLrel. 21; Created)
DT	01-JUN-2002 (TReMBLrel. 21; Last sequence update)
DT	01-MAR-2003 (TReMBLrel. 23; Last annotation update)
DE	Lactotransferrin.
OS	Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96081613; PubMed=8551695;
RA Sato I.;
RT "Characterization of the 84-kDa protein with ABH activity in human
seminal plasma";
RL Jpn. J. Legal Med. 49:281-293(1995).
DR HSP: 077811; IBIX.
DR InterPro: IPR001156; Transferrin.
DR Pfam: PF00405; transferrin; 1
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDBE6 CRC64;

Query Match 54.1%; Score 144.5; DB 4; Length 38;
Best Local Similarity 81.6%; Pred. NO. 4.3e-13;
Matches 31; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 GR-RSVQWCAVSQPEATKCFQWQNMKVRGPPVSCI 37
DB 1 GRXRSVQWCAVSQPEADKXFQWQNMKVRGPPVSCI 38

RESULT 7
Q8WMN8 PRELIMINARY; PRT; 704 AA.
AC Q8WMN8
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Lactoferrin (Fragment).
GN LTF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Wang S.-R., Lin T.-Y., Weng C.-N.;
RT "Isolation and expression of porcine milk lactoferrin.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
DR EMBL: L77987; AAL40161.1;
DR InterPro: IPR001156; Transferrin.
DR Pfam: PF00405; transferrin; 2.
DR PRINTS: PR00422; TRANSFERRIN.
DR SMART: SM00094; TR_FER; 2.
DR PROSITE: PS00205; TRANSFERRIN_1; 2.
DR PROSITE: PS00206; TRANSFERRIN_2; 2.
DR PROSITE: PS00207; TRANSFERRIN_3; 1.
KW Glycoprotein; Iron transport; Metal-binding; Transport.
FT NON_TER 1
SQ SEQUENCE 704 AA; 77681 MW; 64EE769F7503CC32 CRC64;

Query Match 44.9%; Score 120; DB 6; Length 704;
Best Local Similarity 43.5%; Pred. NO. 2.3e-08;
Matches 20; Conservative 10; Mismatches 14; Indels 2; Gaps 1;

QY 3 RRSVQWCAVSQPEATKCFQWQNMKVRGPPVSCIQIOA 48
DB 22 KGVQWCAVSQPEATKCFQWQNMKVRGPPVSCIQIOA 65

RESULT 8
Q9TR80 PRELIMINARY; PRT; 33 AA.
AC Q9TR80
DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Lactoferrin (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=95127729; PubMed=7827104;
RA Qian Z.Y., Jolles P., Migliore-Samouri D., Fiat A.M.;
RL Biochim. Biophys. Acta 1243:25-32(1995).
DR HSP: 077698; ICE2.
DR InterPro: IPR001156; Transferrin.
DR Pfam: PF00405; transferrin; 1.
SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 43.1%; Score 115; DB 6; Length 33;
Best Local Similarity 56.7%; Pred. NO. 5.3e-09;
Matches 17; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 3 RRSVQWCAVSQPEATKCFQWQNMKVRGPP 32
DB 3 RKNVQWCAVSQPEATKCFQWQNMKVRGPP 32

RESULT 9
Q9DBD0 PRELIMINARY; PRT; 700 AA.
AC Q9DBD0
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE 1300017J02Rik protein.
GN 1300017J02Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Saito K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavola T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
DR EMBL: AK005035; BAB23762.1;
DR HSP: P19134; ITFD.
DR MGI: MGI1919025; 1300017J02Rik.
DR InterPro: IPR001156; Transferrin.
DR Pfam: PF00405; transferrin; 2.
DR PRINTS: PR00422; TRANSFERRIN.
DR SMART: SM00094; TR_FER; 2.
DR PROSITE: PS00205; TRANSFERRIN_1; 1.
DR PROSITE: PS00207; TRANSFERRIN_3; 1.

```

RX MEDLINE=84307580; PubMed=6236811;
RA Aldred A.R., Howlett G.J., Schreiber G.;
RT "Synthesis of rat transferrin in Escherichia coli containing a
RL Biochem. Biophys. Res. Commun. 122:960-965(1984).
DR EMBL; M261113; AAA42266.1;
DR HSSP; P02787; 1A8F.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
FT NON_TER 1.
FT NON_TER 87
SQ SEQUENCE 87 AA; 9487 MW; 1982C1D3FDD3BA2 CRC64;

Query Match 34.6%; Score 92.5; DB 11; Length 87;
Best Local Similarity 41.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 11; Mismatches 10; Indels 3; Gaps 1;

QY 4 RSVQCAVSQPEATKCFQWQNRKV---RGPPVSCIQRDS 41
Db ::::::::::: ||| ::::: ||| ::::: |||
17 KTVXWCASEHENTKICISFRDHMKVLPADGPRACVVKTS 57

RESULT 14
Q9PT13 PRELIMINARY; PRT; 691 AA.
AC Q9PT13;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Transferrin.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98079486; PubMed=9418291;
RA Tange N., Jong-Young L., Mikawa N., Hirono I., Aoki T.;
RT "Cloning and characterization of transferrin cDNA and rapid detection
RT of transferrin gene polymorphism in rainbow trout (Oncorhynchus
RT mykiss).";
RL Mol. Mar. Biol. Biotechnol. 6:351-356(1997).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
DR EMBL; D89083; BAA84103.1;
DR HSSP; P56410; 1AOV.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
KW Glycoprotein; Iron transport; Metal-binding; Transport.
SQ SEQUENCE 691 AA; 75084 MW; E3131E0F7AC1D058 CRC64;

Query Match 32.4%; Score 86.5; DB 13; Length 691;
Best Local Similarity 37.2%; Pred. No. 0.0012;
Matches 16; Conservative 8; Mismatches 14; Indels 5; Gaps 1;

QY 6 VQWCAVSQPEATKCFQWQNRKVGGPPVSCIKRDSPIQICQA 48
Db ::::::::::: ||| ::::: ||| ::::: |||
25 VRWCVKSDKELQKCHDLAANVAQ-----FSCVRDSDSLEICQA 62

RESULT 15
Q8UUK7 PRELIMINARY; PRT; 517 AA.
ID Q8UUK7
AC Q8UUK7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
```

```

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Transferrin (Fragment).
GN TF.
OS Oncorhynchus kisutch (Coho salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8019;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-02;
RX MEDLINE=99363267; PubMed=10434423;
RA Ford M.J., Thornton P.J., Park L.K.;
RT "Natural selection promotes divergence of transferrin among salmonid
RT species.";
RL Mol. Ecol. 8:1055-1061(1999).
DR EMBL; AF114866; AAF03083.1;
DR EMBL; AF114855; AAF03083.1; JOINED.
DR EMBL; AF114856; AAF03083.1; JOINED.
DR EMBL; AF114857; AAF03083.1; JOINED.
DR EMBL; AF114858; AAF03083.1; JOINED.
DR EMBL; AF114859; AAF03083.1; JOINED.
DR EMBL; AF114860; AAF03083.1; JOINED.
DR EMBL; AF114861; AAF03083.1; JOINED.
DR EMBL; AF114862; AAF03083.1; JOINED.
DR EMBL; AF114863; AAF03083.1; JOINED.
DR EMBL; AF114864; AAF03083.1; JOINED.
DR EMBL; AF114865; AAF03083.1; JOINED.
DR InterPro; IPR001156; transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 1.
FT NON_TER 1.
FT NON_TER 517
SQ SEQUENCE 517 AA; 56118 MW; 7F99938E11B2205C CRC64;

Query Match 32.0%; Score 85.5; DB 13; Length 517;
Best Local Similarity 41.9%; Pred. No. 0.0012;
Matches 18; Conservative 6; Mismatches 14; Indels 5; Gaps 2;

QY 6 VQWCAVSQPEATKCFQWQNRKVGGPPVSCIKRDSPIQICQA 48
Db ::::::::::: ||| ::::: ||| ::::: |||
14 VRWCVKSEKELKCHDLAAN---VAG--FSCVRDSDSLEICQA 51

Search completed: August 11, 2003, 08:20:53
Job time : 29.9612 secs
```



Creation date: 12-31-2003
Indexing Officer: ANKIPIGI - ALLISON NKPIGI
Team: OIPEScanning
Dossier: 09508095

Legal Date: 08-12-2003

No.	Docode	Number of pages
1	CTNF	13
2	892	1
3	FOR	43

Total number of pages: 57

Remarks:

Order of re-scan issued on